Advancing airborne Doppler lidar wind measurements for atmospheric boundary layer research

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Philipp Gasch, M.Sc., M.A.,
aus Riesa

Tag der mündlichen Prüfung: 27.11.2020
Referent: Prof. Dr. Christoph Kottmeier
Korreferent: Prof. Dr. Markus Rapp

Abstract

There is a need for improvement of observational capabilities with respect to wind measurements in the atmospheric boundary layer (BL). This work is dedicated to the commissioning of a new airborne Doppler lidar (ADL) system, which provides wind measurements inside the turbulent BL at highest resolution up to date. To assess the potential of such measurements and address new challenges presented by them, a two-step approach based on virtual observations and real-world measurements is employed.

As a first part of this work, a novel large eddy simulation (LES)-based airborne Doppler lidar simulator (ADLS) is developed. In the ADLS, simulated measurements are conducted based on LES wind fields, which provides the advantage of knowing the 3-D wind field along the lidar measurement curtain. In a first step, the ADLS is used to assess the error in ADL retrieved wind profiles, due to the violation of the flow homogeneity assumption postulated in the retrieval, for a common system setup and retrieval strategy. In a second step, the ADLS is used to identify preferable system setups and to quantify retrieval strategy limits beyond which ADL wind profiling becomes unreliable. Additionally, the ADLS is used to validate a proposed data-driven uncertainty estimation method, offering new possibilities for wind profile accuracy assessment.

In the second part of this work, real-world measurements obtained on four research flights with a new prototype ADL system are presented. In data processing, the motion correction accuracy, time synchronization, beam alignment and lidar signal quality control is evaluated. It is shown that the measured ground return velocities allow for a flight state dependent assessment of error in the inertial navigation system, which is used for motion correction. A refined ground return based motion correction procedure is demonstrated to be beneficial for measurement accuracy. The results show that reliable measurements are possible even when the aircraft is flying inside the turbulent BL. Thus, in combination with the ADLS results, the accuracy of the retrieved wind profiles is assessed. Further, it is shown that the vertical wind mean and variance, and their associated uncertainties, can be quantified from ADL measurements by taking into account the fixed-beam geometry, lidar pulse averaging effects, as well as random and systematic uncertainties due to the finite measurement lengths. Finally, the analysis of three application studies demonstrates that the prototype system provides substantial new insight into flow and turbulence characteristics of the BL.

Summary

The Earth's ever changing atmosphere requires observations to enable understanding and prediction of processes relevant for weather forecasts and climate projections. The atmospheric boundary layer (BL) is the part of the atmosphere in which everyday life takes place, and which drives exchange processes between the surface and the free troposphere. Hence, the BL is of great relevance for atmospheric science. As such, a core dynamic variable inside the BL is the wind, which is also of general importance for process understanding and model evaluation. Despite its relevance, the BL is still regarded as under-sampled with respect to vertical profiles of basic atmospheric variables such as wind, temperature and humidity. Therefore, the development of new observational capabilities is required. In this respect, the development of airborne Doppler lidar (ADL) systems for high resolution BL wind measurements is deemed important, as they can advance current observational capabilities.

ADL can provide valuable information by allowing for spatially resolved and targeted measurements of atmospheric flow phenomena. Often, ADL systems are used to retrieve spatially resolved vertical profiles of the horizontal wind through the airborne velocity azimuth display technique (AVAD). Another important application are investigations of the vertical wind by means of direct nadir stares. The majority of existing ADL systems are implemented on fast-flying aircraft. Thus, their spatial resolution is coarse (e.g. the most utilized system has an along-track resolution of 5 km for wind profiling and 200 m for nadir measurements), limiting their applicability for BL studies. Only three case studies using high resolution ADL wind profiling systems on slow-flying medium-range turboprop aircraft are available up to date, illustrating the potential and necessity of high resolution ADL measurements for BL research.

This work introduces a new prototype ADL system onboard a slow-flying Dornier 128-6 medium-range turboprop research aircraft (call-sign D-IBUF) to extend the observational capabilities available for BL research and show their added value. High resolution ADL observations in the turbulent BL present new challenges, both for wind profiling and nadir measurements. Hence, this work is structured into two parts. In the first part, to address these challenges and provide a link between theory and measurements, a novel large eddy simulation (LES)-based airborne Doppler lidar simulator (ADLS) is presented. In the ADLS, simulated measurements are conducted based on LES wind fields, considering the coordinate and geometric transformations ap-

plicable to real-world measurements. The ADLS provides added value as the three-dimensional wind field along the lidar measurement curtain is known exactly, which is nearly impossible in real-world situations. Thus, valuable insight can be gained into measurement system characteristics and retrieval strategies. In the second part of this work, real-world ADL measurements from four research flights are analyzed in-depth. An extensive data processing and quality controlling scheme is applied, which quantifies the measurement accuracy of high resolution ADL measurements in the BL. Thereby, in combination with the ADLS results, the retrieval of wind profiles in the turbulent BL is enabled. The retrieved wind profiles facilitate the presentation of spatially resolved nadir measurements of the vertical wind in the BL at unprecedented resolution. Finally, the added value of ADL measurements for BL research is illustrated through three application studies.

In the first part of this work the ADLS enables an in-depth investigation of ADL wind profiling characteristics. For commonly used AVAD retrieval techniques, flow homogeneity is assumed throughout the retrieval volume, a condition which is violated in turbulent BL flow. Assuming an ideal measurement system, the ADLS allows to isolate and evaluate the error in the retrieved wind profiles which occurs due to the violation of the flow homogeneity assumption. Overall, the ADLS demonstrates that wind profiling is possible in turbulent wind fields with errors that allow for meaningful measurements.

In a second step the ADLS is used to investigate the influence of varying system setups and retrieval strategies on wind profiling error. Thereby, it is possible to identify preferable system setups and quantify retrieval strategy limits beyond which ADL wind profiling becomes unreliable. It is shown that quality control of the retrieved wind profiles with the condition number allows for identification of unreliable retrievals, e.g. due to clouds blocking sectors of the scan. Based on these results, an optimized system setup and retrieval strategy is presented, which allows for high resolution and high accuracy wind profiling in the turbulent BL.

Additionally, the ADLS is used to validate a data-driven uncertainty estimation method in a third step. The proposed method consists of a covariance matrix based uncertainty estimation, which uses the residual radial velocity variance, obtained from the least-square fit used for AVAD retrieval, as a proxy for the unknown data variance. Although this procedure has been suggested by theory and previous studies, its applicability has not been vetted systematically before, because knowledge of the input truth is unavailable in real-world ADL measurements. Results show that the retrieval error caused by flow inhomogeneity smaller than the scanned volume can be conservatively estimated based on the measured data through using the effective sample-size corrected residual radial velocity variance. However, the retrieval error due to flow inhomogeneity in the order of and larger than the scanned volume cannot be quantified from the

measurements. Hence, a parameterization of this contribution is developed based on BL and ADL system parameters available in real-world measurements. In combination, the data-driven uncertainty estimation method is shown to yield reliable, conservative results for a wide range of system setups and retrieval strategies under varying atmospheric conditions.

In the second part of this work, real-world measurements obtained with a new prototype ADL system based on four research flights conducted in the vicinity of Brunswick (Germany) in July 2019 are presented. The measurement accuracy of real-world ADL measurements inside the BL is analyzed and quantified. To this end, existing procedures from a number of previous studies are combined, refined and extended, leading to an in-depth data processing and quality controlling scheme. Besides the conducted time synchronization, beam alignment and lidar signal quality control, a focus is put on the evaluation of the motion correction accuracy. It is shown that the measured ground return velocities allow for a flight state dependent assessment of error in the inertial navigation system used for motion correction. A refined ground return based motion correction procedure is introduced and demonstrated to be beneficial for measurement accuracy. The results show that reliable measurements are possible even when the aircraft is flying inside the turbulent BL.

Combining the accuracy estimates and the ADLS developed data-driven uncertainty estimation method, the uncertainty in the ADL retrieved vertical profiles of the horizontal wind is then assessed. Following this, the quality controlled wind profiles are used to enable high resolution nadir retrievals of the vertical wind. The slower aircraft speed and higher lidar measurement frequency provide a one order of magnitude higher along-track resolution than with previously existing systems. Thereby, taking into account the fixed-beam geometry, lidar pulse averaging effects as well as random and systematic uncertainties due to the finite measurement lengths, this work shows for the first time that a quantification of the vertical wind variance and its associated uncertainties is possible from ADL measurements inside the BL.

Finally, the analysis of three application studies demonstrates that the prototype system provides substantial new insight into flow and turbulence characteristics in the BL. In the first application study the prototype ADL is used for a showcase retrieval of vertical wind statistics in a clear section of a turbulent BL topped with shallow cumulus humilis elsewhere. Added value is provided as the ADL is able to retrieve the vertical wind statistics at all levels simultaneously with reasonable uncertainty. In a second application study the potential of the prototype ADL system to provide spatially resolved measurements of both the horizontal and vertical wind in moderately complex terrain is utilized. Thereby, both the temporal and spatial development of the BL as well as its link to meso-scale orographic features can be captured. A third application

study analyzes differences in BL turbulence characteristics observed between across- and along-wind sampling.

In conclusion, this work highlights the potential and importance of high resolution ADL measurements for extending the observational capabilities available for BL research. A new contribution to the theoretical study of ADL systems is introduced through the ADLS, which allows for previously inaccessible insight into the wind profiling retrieval process. The real-world ADL measurements at highest resolution up to date present a step towards flexible and accurate measurements of flow phenomena in the atmospheric BL under a wide range of conditions.

Zusammenfassung

Die ständigem Wandel unterliegende Erdatmosphäre erfordert Beobachtungen, um die für Wettervorhersagen und Klimaprojektionen relevanten Prozesse zu verstehen. In der Erdatmosphäre ist die atmosphärische Grenzschicht (BL) von großer Bedeutung, da sie Austauschprozesse zwischen der Oberfläche und der freien Troposphäre antreibt und da sich das tägliche Leben in ihr abspielt. Folglich ist das BL-Windfeld eine wichtige dynamische Größe und damit von Bedeutung für Prozessverständnis und Modellevaluation in der Atmosphärenforschung. Trotz ihrer Relevanz kann die BL hinsichtlich der vertikalen Profile grundlegender atmosphärischer Größen wie Wind, Temperatur und Feuchte immer noch nicht ausreichend durch Messungen erfasst werden. Daher ist die Entwicklung neuer Beobachtungssysteme erforderlich. In dieser Hinsicht stellt die Entwicklung von flugzeuggetragenen Doppler-Lidar Systemen (ADL) für hochauflösende BL-Windfeldmessungen einen wichtigen Baustein für die Erweiterung der vorhandenen Beobachtungskapazitäten dar.

ADL kann wertvolle Informationen liefern, indem es räumlich aufgelöste und gezielte Messungen atmosphärischer Strömungsphänomene ermöglicht. Häufig werden ADL-Systeme verwendet, um räumlich aufgelöste Vertikalprofile des horizontalen Winds durch die flugzeuggetragene Geschwindigkeits-Azimut-Methode (AVAD) zu erhalten. Eine weitere wichtige Anwendung sind Untersuchungen des Vertikalwinds mittels direkter Nadir-Messungen. Mit Ausnahme von zwei Fallstudien, die das Potenzial von hochauflösenden ADL-Windprofil-Messungen auf langsam fliegenden Turboprop-Flugzeugen veranschaulichen, sind alle existierenden ADL-Systeme auf schnell fliegenden Flugzeugen installiert. Daher besitzen sie nur eine grobe räumliche Auflösung für Windprofil- und Nadir-Messungen, was ihre Anwendbarkeit für BL-Studien einschränkt.

Diese Arbeit stellt den Prototyp eines neuen ADL-Systems an Bord eines langsam fliegenden Turboprop-Forschungsflugzeugs (Dornier 128-6, Rufzeichen D-IBUF) vor, um die für die BL-Forschung verfügbaren Beobachtungskapazitäten zu erweitern und den Mehrwert des neuen Messsystems aufzuzeigen. Hochauflösende ADL-Beobachtungen in der turbulenten BL stellen neue Herausforderungen dar, sowohl für Windprofil- als auch für Nadir-Messungen. Daher ist diese Arbeit in zwei Teile gegliedert. Um die Herausforderungen zu überwinden und eine Verbindung zwischen Theorie und Messungen herzustellen wird im ersten Teil ein neuarti-

ger LES (Grobstruktursimulations)-basierter Flugzeug-Doppler-Lidar-Simulator (ADLS) vorgestellt. Im ADLS werden simulierte Messungen auf der Grundlage von LES-Windfeldern durchgeführt, wobei alle in der Realität anzuwendenden Koordinatensystem- und geometrischen Transformationen berücksichtigt werden. Der ADLS erlaubt neue Einblicke, da die zur Simulation der Messungen verwendeten Windfelder exakt bekannt sind, was in realen Messsituationen nahezu unmöglich ist. Auf diese Weise kann wertvolles Verständnis über die Eigenschaften des Messsystems und die Messmethoden gewonnen werden. Im zweiten Teil dieser Arbeit werden reale ADL-Messungen aus vier Forschungsflügen eingehend analysiert. Es wird ein umfangreiches Datenverarbeitungs- und Qualitätskontrollschema angewandt, das die Messgenauigkeit von hochauflösenden ADL-Messungen in der BL quantifiziert. Dadurch wird in Kombination mit den ADLS-Ergebnissen die Messung von Windprofilen in der turbulenten BL ermöglicht. Die gemessenen Windprofile wiederum ermöglichen die Prozessierung von räumlich aufgelösten Nadir-Messungen des Vertikalwinds in der BL mit bisher unerreichter Auflösung. Schließlich wird der Mehrwert der neuen ADL-Messungen für die BL-Forschung durch drei Anwendungsstudien veranschaulicht.

Im ersten Teil dieser Arbeit ermöglicht der ADLS eine genaue Untersuchung der Eigenschaften von ADL-Messverfahren zur Gewinnung von Windprofilen mittels AVAD. Häufig angewendete AVAD-Messverfahren setzen Homogenität der Strömung im Messvolumen voraus. Diese Bedingung ist in turbulenter BL-Strömung verletzt. Unter der Annahme eines idealen Messsystems ermöglicht der ADLS die Isolierung und Bewertung des Fehlers, der bei der Messung von Windprofilen mittels AVAD aufgrund der Verletzung der Annahme der Strömungshomogenität verursacht wird. Insgesamt zeigt der ADLS, dass die Gewinnung von Windprofilen in turbulenten Windfeldern mit akzeptablen Fehlern behaftet ist die aussagekräftige Messungen ermöglichen.

In einem zweiten Schritt wird der ADLS verwendet um den Einfluss unterschiedlicher Systemkonfigurationen und Messverfahren auf den Fehler in den gewonnenen Windprofilen zu untersuchen. Dadurch ist es möglich, bevorzugte Systemkonfigurationen zu identifizieren und die Grenzen der Messmethoden zu erkennen, jenseits derer die Gewinnung von BL-Windprofilen unzuverlässig wird. Es wird gezeigt, dass die Qualitätskontrolle der generierten Windprofile mit der Konditionszahl die Identifizierung unzuverlässiger Windprofile ermöglicht. Basierend auf diesen Ergebnissen wird eine optimierte Systemkonfiguration und angepasstes Messverfahren vorgestellt, welches die Gewinnung von Windprofilen mit hoher Auflösung und Genauigkeit auch in der turbulenten BL erlaubt.

Zusätzlich wird der ADLS in einem dritten Schritt zur Validierung einer datenbasierten Unsicherheitsschätzmethode verwendet. Die vorgeschlagene Methode nutzt eine auf der Kovari-

anzmatrix basierende Unsicherheitsschätzung, bei der die Residual-Varianz der Radialgeschwindigkeit aus der AVAD-Messanpassung (Methode der kleinsten Quadrate) als Proxy für die unbekannte Datenvarianz verwendet wird. Obwohl dieses Verfahren aus der Theorie bekannt ist und in vorhergehenden Studien vorgeschlagen wurde, konnte seine Anwendbarkeit bisher noch nicht systematisch überprüft werden, da die Kenntnis der Eingangswahrheit bei realen ADL-Messungen nahezu unmöglich ist. Die Ergebnisse zeigen, dass der Fehler im gewonnenen Windprofil, der durch Strömungsinhomogenität auf Skalen kleiner als das Messvolumen verursacht wird, anhand der gemessenen Daten konservativ geschätzt werden kann. Dazu muss die Residual-Varianz der Radialgeschwindigkeit aus der AVAD-Messanpassung für die effektive Probengröße korrigiert werden. Der Fehler im generierten Windprofil aufgrund von Strömungsinhomogenität auf der Skala des gescannten Volumens und größer kann jedoch aus den Messungen nicht quantifiziert werden. Daher wird eine Parametrisierung dieses Beitrags auf der Grundlage von BL- und ADL-Systemparametern entwickelt, die in realen Messungen verfügbar sind. In Kombination wird gezeigt, dass die datenbasierte Unsicherheitsschätzmethode zuverlässige, konservative Ergebnisse für ein breites Spektrum von Systemkonfigurationen und Messverfahren bei unterschiedlichen atmosphärischen Bedingungen liefert.

Im zweiten Teil dieser Arbeit werden reale Messungen vorgestellt, die mit einem neuen Prototyp ADL-System während vier Forschungsflügen in der Nähe von Braunschweig (Deutschland) im Juli 2019 durchgeführt wurden. Die Messgenauigkeit von realen ADL-Messungen innerhalb der BL wird analysiert und quantifiziert. Zu diesem Zweck werden bestehende Methoden aus einer Reihe von früheren Studien zusammengeführt und erweitert, die somit eine umfassende Datenverarbeitung und Qualitätskontrolle erlauben. Neben der durchgeführten Zeitsynchronisation, der Ermittelung der Strahlausrichtung und der Qualitätskontrolle des Lidarsignals wird ein Schwerpunkt auf die Bewertung der Genauigkeit der Bewegungskorrektur gelegt. Es wird gezeigt, dass die mittels ADL gemessenen Bodengeschwindigkeiten eine Bewertung der Fehler des für die Bewegungskorrektur verwendeten Trägheitsnavigationssystems in Abhängigkeit des Flugzustands ermöglichen. Ein erweitertes Bewegungskorrekturverfahren wird vorgestellt und als vorteilhaft für die Messgenauigkeit demonstriert. Die Ergebnisse zeigen, dass zuverlässige Messungen auch dann möglich sind, wenn das Flugzeug innerhalb der turbulenten BL fliegt.

Durch Kombination der Genauigkeitsschätzungen und der im ADLS entwickelten datengetriebenen Unsicherheitsschätzmethode wird sodann die Unsicherheit in den mittels ADL gewonnenen horizontalen Windprofilen bewertet. Im Anschluss daran werden die qualitätskontrollierten Windprofile verwendet, um hochauflösende Nadir-Messungen des Vertikalwinds zu ermöglichen. Im Vergleich zu bisher existierenden Systemen wird dabei aufgrund der langsameren Fluggeschwindigkeit und höheren Lidar-Messfrequenz eine Erhöhung der Auflösung (in Flugrich-

tung) um eine Größenordnung erreicht. Dadurch ist unter Berücksichtigung der unbeweglichen Strahlausrichtung, der Lidar-Mittelungseffekte im Messvolumen sowie der zufälligen und systematischen Unsicherheiten aufgrund der endlichen Messlängen eine Quantifizierung der Varianz des Vertikalwinds und der damit verbundenen Unsicherheiten möglich.
Schließlich zeigt die Analyse von drei Anwendungsstudien, dass das Prototyp ADL-System wesentliche und neue Informationen über Strömungs- und Turbulenzeigenschaften der BL liefert. In der ersten Anwendungsstudie wird der ADL-Prototyp für eine Studie des Vertikalwinds in einem wolkenlosen Abschnitt einer andernorts mit flachen Cumulus humilis bedeckten BL verwendet, Der Mehrwert des ADL-Prototyps besteht in der gleichzeitigen Verfügbarkeit der statistischen Kenngrößen des Vertikalwinds auf allen Ebenen, so z.B. der Varianz, bei akzeptabler Unsicherheit. In einer zweiten Anwendungsstudie wird das Potential des Prototyp ADL-Systems genutzt, um räumlich aufgelöste Windmessungen des Horizontal- als auch des Vertikalwinds in mäßig komplexem Gelände zu liefern. Dadurch können sowohl die zeitliche und räumliche Entwicklung der BL als auch deren Verknüpfung mit meso-skaligen orographischen Merkmalen erfasst werden. Eine dritte Anwendungsstudie analysiert Unterschiede in den beobachteten BL-Turbulenzcharakteristika für Flüge quer und längs zur vorherrschenden Windrichtung.
Abschließend unterstreicht diese Arbeit das Potenzial und die Bedeutung hochauflösender ADL-Messungen für die Erweiterung der Beobachtungskapazitäten in der BL-Forschung. Ein neuer Beitrag zur theoretischen Untersuchung von ADL-Systemen wird durch den ADLS eingeführt, der bisher unzugängliche Einblicke in den Messprozess ermöglicht. Die realen ADL-Messungen mit höchster Auflösung stellen einen Schritt in Richtung flexibler und präziser Messungen von Strömungsphänomenen in der atmosphärischen BL unter einer großen Bandbreite von Bedingungen dar.

Contents

1. Introduction [1]

The atmospheric boundary layer (BL) is of great relevance for atmospheric science as it provides the stage for exchange processes between the surface and the free troposphere. The BL state is challenging to adequately capture in observations and represent in numerical modeling because the turbulent processes present cover a wide range of scales. Thus, to facilitate process understanding and model evaluation, high resolution observations of the BL are important. The wind is a core dynamical variable in the atmosphere. Flow in the atmosphere can exhibit strong spatial and temporal variability and is thereby challenging to measure, a problem which is intensified further in the BL where flow is generally turbulent. Despite the great variability and importance, the majority of wind observations conducted by operational networks and for research purpose are limited to the surface and therefore do not capture the vertical structure of the wind field adequately. Additionally, almost all observations are conducted as point measurements and hence struggle to capture the spatial variation of atmospheric flows.

Operational weather observation strongly relies on radiosondes for the routine observation of vertical profiles of the horizontal wind (also called wind profiles in the following). Additional wind information is provided by aircraft measurements on-board regular passenger flights, satellite-based atmospheric motion vectors from cloud tracking (Deutscher Wetterdienst, 2020b). Despite the various observation systems, up to date the atmospheric wind field is regarded as strongly under-sampled, especially in the vertical dimension (Baker et al., 2014, and references therein). This under-sampling is even more severe with respect to the BL, where very localized flow phenomena and turbulent flow are present. To capture the spatial heterogeneity and high frequency changes of BL flow, measurements at highest resolution are necessary, requiring the development of new measurement capabilities (Emeis et al., 2018; Geerts et al., 2018).

Throughout the last decades, Doppler lidar technology has emerged which enables active remote sensing of atmospheric winds in clear air (Weitkamp, 2005). Today, the Doppler lidar technology has achieved maturity and the integration of ground-based Doppler lidar into ob-

[1] The introduction is adapted from the publication: Gasch, P., Wieser, A., Lundquist, J. K. and Kalthoff, N., 2020: An LES-based airborne Doppler lidar simulator and its application to wind profiling in inhomogeneous flow conditions. Atmospheric Measurement Techniques, 13(3), 1609-1631.

servational networks is beginning. For research purposes, multiple Doppler lidar systems have been combined to allow for a spatially resolved measurement of atmospheric flow at fixed locations (Träumner et al., 2015; Fernando et al., 2019; Adler et al., 2020a). Recently, the first satellite-based Doppler lidar has been deployed to space (European Space Agency, 2018). Due to its unique capabilities, its data is already being assimilated into operational weather models (Deutscher Wetterdienst, 2020c), while still undergoing validation (Witschas et al., 2020).
Airborne Doppler lidar (ADL) systems can provided important additional information for the measurement of flow phenomena by enabling spatially resolved (De Wekker et al., 2012; Chouza et al., 2016a; Witschas et al., 2017), as well as targeted observations with rapid deployment (Weissmann et al., 2005b; Zhang et al., 2018). Further, ADL 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; Baidar et al., 2018; Tucker et al., 2018; Witschas et al., 2020).
The most utilized task of ADL systems is spatially resolved wind profiling (Weissmann et al., 2005b; De Wekker et al., 2012; Bucci et al., 2018). Wind profiling describes the indirect measurement (also called retrieval, the applied technique is discussed below) of wind profiles in retrieval volumes along the flight path. Further, some ADL systems also allow for observation of the vertical wind by means of nadir stares (Kiemle et al., 2007, 2011; Chouza et al., 2016b; Witschas et al., 2017). The development of new observational capabilities, especially with respect to ADL, has been identified as a cornerstone for the improvement of current observational capabilities to adequately capture the BL state (Geerts et al., 2018).
Up to date, most existing ADL systems are based on fast (and typically high) flying aircraft. Due to the fast aircraft speed, the spatial resolution of the measurements is challenging for BL studies. For example, the most utilized ADL system based on a Falcon 20 aircraft is able to retrieve a wind profile every 6 km along the flight path (Weissmann et al., 2005b; Witschas et al., 2017). This resolution is very coarse for BL studies, where flow phenomena on much smaller scales are present. The same ADL system also allows for measurements of the vertical wind by means of nadir stares with a lidar measurement frequency of 1 Hz. Thus, at a typical aircraft speed of $200\,\mathrm{m\,s^{-1}}$, measurements of the vertical wind can be obtained every 200 m (Kiemle et al., 2011; Chouza et al., 2016b; Witschas et al., 2017), which is also coarse for BL studies. Besides the coarse resolution, limitations also exist in the applicability of fast aircraft for measurements inside the BL due to airspace restrictions, as their operation under visual flight rules (VFR) is often problematic. Nevertheless, two studies have been conducted using the above system (Kiemle et al., 2007, 2011), with promising results motivating further application of ADL for BL studies.

The only reported BL studies of ADL measurements based on slow (and typically low) flying aircraft are given by De Wekker et al. (2012), Godwin et al. (2012) and Schroeder et al. (2020). The ADL system presented by De Wekker et al. (2012) and Godwin et al. (2012) is limited to wind profiling with a resolution of $1.5-2\,\mathrm{km}$ along the flight path. Hence, meso-scale flow phenomena (e.g. flow across the Salinas valley in California) are investigated in a case study. The study by Schroeder et al. (2020) does present wind profiling and nadir measurement geometries for flights around a wildfire. However, the study lacks an analysis on the achieved measurement accuracy as well as the potential of the measurements, as it is focused on the technical system description and thereby only provides quicklooks of the acquired data. Hence, up to date, an investigation of the potential for high resolution analysis of the vertical wind inside the BL by means of nadir staring ADL is missing.

Most currently used airborne wind profiling approaches apply the airborne velocity azimuth display technique (AVAD) or closely related retrieval techniques to estimate vertical profiles of the horizontal wind in the retrieval volume (Leon and Vali, 1998). Therein, scanning in conical circles along the flight path is conducted at a constant off-nadir angle. During each scan, a portion of the horizontal wind is projected into the radial velocity measurement in dependence of the azimuth viewing direction. Thus, the azimuth variation of the radial velocity can be used to infer the underlying average horizontal wind. AVAD approaches assume homogeneous flow conditions throughout the retrieval volume (e.g. no turbulent deviations from the mean flow). Especially in turbulent environments such as the BL, this flow homogeneity assumption is rarely fulfilled, which introduces error in the retrieved wind profiles. Consequently, 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 wind profiling systems use high scan elevation angles (closer to nadir). Reasons for doing so include constraint of the measurement footprint, hardware and range limitations or signal attenuation. As a result, the measured radial velocity is strongly influenced by variations of the vertical wind along the scan circle. Thereby, the effect of violations of the homogeneous flow assumption is intensified, causing in non-negligible errors in the retrieved wind profiles (Tian et al., 2015; Bucci et al., 2018).

A common method to assess the reliability of AVAD retrieved wind profiles is by comparison to wind profiles obtained using other measurement systems, most commonly dropsondes released from the same aircraft. However, this comparison is expensive to conduct and suffers from co-location problems of the measurements (Weissmann et al., 2005b; Witschas et al., 2020, and references therein). This difference often makes a direct comparison of the results challenging, as it is difficult to determine if the observed errors occur due to the differing sampling volumes

and characteristics, violation of the flow homogeneity assumption or other instrument errors. The discussed problems are worsened in turbulent BL flow, where rapid changes in flow occur on small scales and the violation of the flow homogeneity assumption is expected to be non-negligible.

Due to these challenges, theoretical study through idealized simulations of Doppler measurement systems can be useful. Idealized simulations based on known input wind fields provide detailed insight into the capabilities and limitations of the airborne measurement and retrieval process, they have progressed with the availability of more computer power for higher resolution over the years (Schröter et al., 2000; Frehlich, 2001; Sühring and Raasch, 2013). Reitebuch et al. (2001), Leike et al. (2001), Lorsolo et al. (2013) and Guimond et al. (2014) show the importance of model-based simulator studies for airborne Doppler systems, while relying on coarser resolution model output and focused on errors introduced due to the measurement system errors. For a simulation-based investigation of the wind profiling retrieval process in the turbulent BL, high resolution atmospheric wind fields (O(100 m)) are required in order to represent BL turbulence characteristics adequately. Despite the progress, a simulation of an airborne wind profiling system with complex scanning geometries, based on high resolution atmospheric wind fields, was missing before the work presented here.

This work is dedicated to the commissioning of a new ADL system which extends the observational capabilities available for BL research. The new ADL system is implemented on the slow-flying, medium-range turboprop Dornier 128-6 research aircraft (call-sign D-IBUF). As outlined above, high resolution ADL measurements in the turbulent BL environment pose a number of previously unaddressed challenges, both for wind profiling and nadir measurements. Therefore, this work evolves around the question whether accurate and meaningful ADL measurements are possible inside the BL using the new system, despite the challenges presented by the turbulent wind field. More specifically, the research objectives addressed with respect to wind profiling at highest resolution inside the turbulent BL flow in the first part of this thesis are:

1. To determine the impact of the violated flow homogeneity assumption used in AVAD on wind profiling error.
2. To find preferable system setups and retrieval strategies which minimize wind profiling error for measurements at highest resolution.
3. To evaluate whether a data-driven uncertainty assessment of the retrieved wind profiles is possible.

In order to address these problems, a novel large eddy simulation (LES)-based airborne Doppler lidar simulator (ADLS) is introduced by Gasch et al. (2020) as a part of this work. The ADLS enables analysis of AVAD wind profiling error through simulation of an idealized ADL measurement system. In the ADLS, distillation of the wind profiling error due to violation of the AVAD flow homogeneity assumption in the turbulent BL is possible, because the three-dimensional wind field along the lidar measurement trajectory (also called input truth in the following) is known exactly, which is near-impossible in real-world measurements.
In the second part of this thesis, real-world BL measurements conducted with a prototype ADL system are analyzed. The research objectives associated with the commissioning of this new system are as follows:

4. To develop a data processing and quality controlling scheme which provides insight into the achieved measurement platform accuracy.
5. To obtain accuracy estimates both for retrieved wind profiles and nadir measurements.
6. To demonstrate the added value of the new high resolution measurement system through a number of showcase application studies.

This work employs a distinction between the terms error, uncertainty and accuracy based on ISO 5725-1:1994 and JCGM 100:2008, which is detailed in Appendix A.1.
Alongside these research objectives, this work is organized as follows: In Chap. 2 existing ADL systems are presented, the new system is introduced and existing studies dealing with the accuracy estimation of Doppler lidar retrieved wind profiles are discussed. Chap. 3 presents the theory underlying ADL wind measurements as well as an associated proposed uncertainty estimation method for AVAD wind profiling. The chapter ends with a description of LES theory, which poses the foundation of the ADLS. The ADLS and its operation are introduced in Chap. 4, further, the simulation strategy for statistical analysis of wind profiling error is outlined. In Chap. 5 statistical results on wind profiling error obtained using the ADLS are discussed, a system setup and retrieval strategy optimization is performed and the applicability of the proposed data-driven uncertainty estimation method is vetted. In Chap. 6 the measurements obtained with the prototype ADL system are analyzed and a thorough data processing and quality control scheme is developed. Then, the measurement system accuracy is assessed and three application studies are conducted to show the added value of the new ADL system. Finally, Chap. 7 provides the conclusions and gives recommendations for future work.

2. Remote sensing of atmospheric boundary layer winds by means of airborne Doppler lidar

The atmosphere is a gaseous fluid which surrounds the planet Earth. As such, the atmosphere exhibits distinct layering, characterized by changes in the gradient of the vertical temperature profile (Seinfeld and Pandis, 2006). An ubiquitous layer inside the troposphere is the BL. The BL can be characterized as the layer of the atmosphere which is directly influenced by processes at the surface on time scales of about an hour or less (Stull, 1988). It provides exchange between the surface below and the free troposphere above. The BL is also the place where everyday life take places, emissions occur and through which the impact of the weather and climate is felt by society. Thus, its importance in atmospheric science cannot be understated.

The BL is often capped by a distinct temperature inversion, separating it from the flow of the free troposphere above (e.g. flow free from the influence of surface friction). The BL height z_i is an important parameter in the characterization of the BL. However, it is not always easy to identify and many definitions for its determination exist (e.g. based on the temperature profile, sensible heat flux profile or turbulence state, Träumner, 2010). Typical BL depths over land range from a few hundred meters to a few thousand meters, while it is considerably more shallow over oceans (Stull, 1988). Under the presence of synoptic high pressure systems the BL above land features a distinct diurnal cycle. Typically, after sunrise a rapid growth of the BL through the residual layer from the previous night is observed, with a slower subsequent growth during the rest of the day. Towards the night, the net radiation balance of the surface becomes negative and a shallow nocturnal BL forms below a strong nocturnal inversion, with the residual layer above (see Stull, 1988, for an illustration). In contrast to the diurnal cycle in high-pressure systems, in the ascending air-mass regions of a low pressure system or in thunderstorms it is not easy to identify or clearly state the extent of the BL. Often, the cloud bottom is defined as the BL height in such situations. The BL possesses an internal structure which can be further distinguished. Close to the ground a surface layer exists (O(10 m)) in which fluxes are within 10% of their surface value and in which often a super-adiabatic temperature profile is present (Stull, 1988; Kraus, 2008). The surface layer is where most ground-based measurements are conducted. Above the surface layer comes the mixed layer which exhibits a constant potential temperature with height, distinct turbulent flux profiles and still a sub-geostrophic wind speed

according to idealized theory. The mixed layer is capped by an inversion of varying intensity, through which exchange processes with the free troposphere occur by entrainment and detrainment (see Träumner, 2010, for an overview). It should be noted that adjacent to the ground a very thin (O(1 mm)) viscous sublayer exists in which transport occurs through molecular processes (Kraus, 2008).

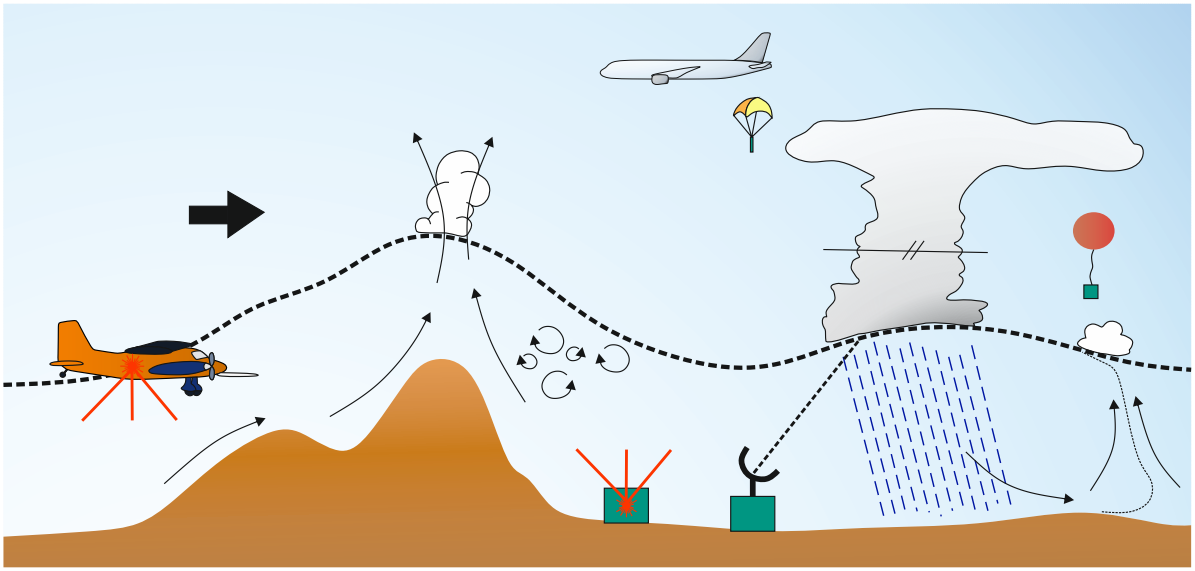

Fig. 2.1.: Schematic illustration of BL variability and measurement systems capable of vertical profiling. Figure not to scale.

The BL height and structure can exhibit strong spatio-temporal variability (Fig. 2.1). For example, in mountainous terrain the BL height is not uniform but follows the terrain instead (De Wekker and Kossmann, 2015; Pal et al., 2016; Lehner and Rotach, 2018). Further, multi-scale advective processes by thermal wind systems play an important role besides the usual convective transport inside the mountainous BL (Weissmann et al., 2005b). Rotor circulations, hydraulic jumps and gravity waves can add further complexity to the flow field (Strauss et al., 2015; Kalthoff et al., 2020). The BL over flat terrain is also not necessarily homogeneous, e.g. due to the super-position of meso-scale flow features (Maurer et al., 2016). Further, intense localized flow features such as thunderstorms can lead to a rapidly changing and heterogeneous BL structure. Prominent examples are cold-pool outflows from meso-scale convective systems, which can travel over hundreds of kilometers and are sometimes associated adverse weather effects such as dust storms (Gasch et al., 2017), besides their high wind speeds (Kuchera and Parker, 2006) and role in convective organization (Tompkins, 2001).

This BL heterogeneity can make ground-based measurements challenging due to representativeness problems, with large observed differences between close-by stations (Maurer et al.,

2016). Further, the reliance of ground-based measurements on the background wind speed for advection often makes the characterization of turbulence problematic (Adler et al., 2019). In summary, the BL complexity and heterogeneity creates a strong need for spatially resolved and targeted observations, especially with respect to the three-dimensional flow field.

2.1. Flow in the atmospheric boundary layer

Processes associated with three-dimensional flow (also called wind, represented by the vector $\boldsymbol{v} = [u, v, w]$ or $u_i = [u_1, u_2, u_3] = [u, v, w]$ in component wise notation) in the atmosphere are described by the Navier-Stokes (NVS) equations (Etling, 2008, besides others):

$$\frac{\partial u_i}{\partial t} + u_k \frac{\partial u_i}{\partial x_k} = -\frac{1}{\rho_a}\frac{\partial p}{\partial x_i} - \delta_{i3} g - \varepsilon_{ijk} f_j u_k + \nu \frac{\partial^2 u_i}{\partial x_k^2}. \tag{2.1}$$

In the above, p is the pressure, ρ_a is the density of dry air and g is the gravitational acceleration. Further, $f_i = [0, 2\Omega_g \cos(\phi_g), 2\Omega_g \sin(\phi_g)]$ is the Coriolis parameter with the Earth's angular speed Ω_g and the geographic latitude ϕ_g and ν is the kinematic viscosity coefficient. Summation over repeated indices is implied, $\delta_{i,3}$ is the Kronecker delta and ε_{ijk} is the Levi-Civita symbol. The left-hand side terms in the NVS equations describe the local change of momentum and its advection by the mean wind. The right-hand side terms describe the effects of the pressure gradient, gravitation, Coriolis-effect and viscosity causing friction. While the NVS equations present a complete description of the atmospheric flow they are not analytically solvable up to date.

A distinct difference between the flow in the free troposphere and the BL exists: Whereas flow in the free atmosphere is typically laminar (except during localized events such as thunderstorms or gravity wave breaking), flow in the BL is turbulent. In contrast to laminar flow, turbulent flow describes a flow state in which three-dimensional vortices on a range of scales are superimposed on the background flow. Turbulent motion is chaotic and cannot be forecast over extended periods or domains. Chaotic in this respect means that while the NVS equations are deterministic, the flow is nevertheless unpredictable (Lorenz, 1963). This behavior is caused by the non-linear term (e.g. products of velocities) in the equations, which leads to a susceptibility to small disturbances in the initial conditions and associated rapid error growth in forecasting.

Reynolds averaging

The NVS equations are valid in all flow conditions. However, their numerical treatment requires a consideration of the relevant process scales, as turbulent processes are too small to be

resolved by coarse numerical weather prediction (NWP) models (whose grid cannot be refined indefinitely, discussed in Sec. 3.7). In order to enable numerical treatment of the NVS equations different simplifying approaches based on scale analysis are used (Etling, 2008). A common approach to introduce a scale separation into the NVS equations is the so-called Reynolds averaging (others are discussed in Sec. 3.7). Reynolds averaging splits the variables in the NVS equations into averages and deviations therefrom (e.g. the averaged grid-scale quantities can be resolved by NWP models, whereas the turbulent fluctuations cannot). As an example, for the vertical wind $w = u_3$, this reads as:

$$w = \overline{w} + w', \tag{2.2}$$

where the overline denotes an average and the prime denotes fluctuations. The appropriate choice of an averaging length is discussed below. Applying the Reynolds averaging to all terms in the NVS equations results in (Etling, 2008):

$$\frac{\partial \overline{u}_i}{\partial t} + \overline{u}_k \frac{\partial \overline{u}_i}{\partial x_k} = -\frac{1}{\overline{\rho_a}} \frac{\partial \overline{p}}{\partial x_i} - \delta_{i3} g - \varepsilon_{ijk} f_j \overline{u}_k + \nu \frac{\partial^2 \overline{u}_i}{\partial x_k^2} - \frac{\partial \overline{u'_k u'_i}}{\partial x_k}. \tag{2.3}$$

Thus, Reynolds averaging of the non-linear advection term introduces a new term which contains the covariance of the velocity fluctuations. In analogy to the viscous friction, the covariance term represents a friction introduced due to turbulence. However, its contribution is orders of magnitude larger than the viscous friction term, therefore it needs to be represented adequately in models. Because of the inability to simulate the turbulent fluctuations directly in NWP, its representation is accomplished through parametrization schemes. Parametrization schemes attempt to represent the covariance contributions (unresolved subgrid-scale processes, e.g. the transfer of momentum, heat and moisture due to turbulence) through relating them to resolved grid-scale variables. To achieve parametrization, a description of turbulence based on bulk properties is employed, characterizing its overall transport efficiency and effect on the atmospheric flow field and energy budget. A short introduction to the properties of turbulence and their measurement is provided in the following to enable a meaningful description and identify relevant bulk properties.

Properties of turbulence

Turbulence accomplishes the transport of mass, momentum, latent and sensible heat as well as other scalar quantities (e.g. aerosols or trace gases) through non-random mixing in the presence of gradients, termed turbulent flux in the following (Stull, 1988; Etling, 2008). Flow in the BL is turbulent due to the high transport efficiency needed for energy exchange processes in the BL.

Due to the mixing, which induces drag, turbulence constantly dissipates energy into heat and thus requires a constant energy supply for its maintenance. The generation of turbulence can be distinguished by its generating force. Mechanical turbulence is generated in stable stratified flow, e.g. due to gradients of the wind speed. A prominent example is the night-time BL where mechanical turbulence is generated close to the surface (as momentum is transported from the atmosphere to the surface). Convective turbulence is present if density differences of the air cause buoyancy forces. A prominent example is the day-time BL in summer, where energy transport processes at the surface create buoyancy forces through temperature and humidity fluctuations. Of course, co-existence of the two turbulence generating mechanisms is the usual case.

Characterization of turbulence

Because of the inability to predict (and measure) turbulent flow in all its three-dimensional detail, current description and prediction resorts to the characterization of the turbulent flow through bulk statistical parameters. These parameters allow for a characterization of the flow, e.g. its ability to transport energy, without full three-dimensional knowledge about the flow present. Thus, the description of turbulence can be simplified from a four dimensional problem to a one or two dimensional problem. An important parameter characterizing turbulent flow is the turbulent kinetic energy (TKE), given as (Stull, 1988):

$$\overline{e} = \frac{1}{2}\left(\overline{u'^2} + \overline{v'^2} + \overline{w'^2}\right). \tag{2.4}$$

As before, primes denote velocity fluctuations and overlines denote averages. Two things become clear from this formulation: First, the process of calculating bulk statistical properties requires averaging over a time or space interval. Second, any observation must allow for high enough resolution to resolve the turbulent deviations in detail. The specification of an adequate averaging time interval and required measurement resolution are discussed below.
An important contribution to the TKE is given through the variance of the vertical wind $\overline{w'^2}$. Thus, insight into the TKE and thereby turbulence characteristics can be gained through studying the vertical wind variance and its associated spectrum, which is therefore an important parameter in BL research. Further, if turbulence is isotropic this implies that $\overline{u'^2} = \overline{v'^2} = \overline{w'^2}$ and thereby $\overline{e} = \frac{3}{2}\overline{w'^2}$. However, BL turbulence can only be assumed to be isotropic at small scales (O(10 m), termed 'local isotropy'), as the presence of the ground and capping inversion lead to a non-random orientation of the vortices for larger scales (Kaimal et al., 1972). Importance is also bestowed on the vertical wind variance through its contribution in the methods

used to parameterize turbulent fluxes. In these methods, fluxes are calculated through combination of the vertical wind variance with the gradient of a scalar component as a proxy for the unknown covariance, e.g. the water vapor mixing ratio and potential temperature for latent and sensible heat, respectively (Kraus, 2008). Similarly, direct measurements of the turbulent fluxes using the eddy-covariance method require the highly resolved vertical wind fluctuations w' as a contribution to the covariance (Kiemle et al., 2007). Due to it's importance, the vertical wind and its variance have been studied for decades. Early studies use aircraft observations (Kaimal et al., 1976; Lenschow et al., 1980, besides others), while later Doppler radar (Kropfli, 1986; Eymard and Weill, 1988, besides others) and Doppler lidar (Eberhard et al., 1989; Lenschow et al., 2000; Hogan et al., 2009; Lothon et al., 2009, besides others) became available.

The statistical bulk parameters describing turbulence can only be calculated with the assumption of horizontally homogeneous turbulence (for aircraft measurements, stationary turbulence for ground-based measurements) and Taylor's hypothesis (mathematical insight into their determination from ADL measurements is provided in Sec. 3.4). Homogeneous turbulence describes the invariance of the turbulence properties to spatial translation. It must be noted that the homogeneous turbulence referred to here is to be distinguished from the frequently used homogeneous/inhomogeneous flow (conditions) used with respect to wind profiling: inhomogeneous flow with respect to wind profiling simply refers to the presence of turbulence, not the inhomogeneity of the turbulence. If homogeneous turbulence is addressed it is specifically referred to as such in this work. Also note that homogeneous turbulence is only assumed in the horizontal, as it is never present in the BL with respect to the vertical due to the presence of the ground and capping inversion.

The assumption of Taylor's hypothesis for the calculation of statistical properties implies that the characteristics of the turbulent vortices do not change while they are being sampled (it is also called 'frozen turbulence' assumption therefore). This assumption is fulfilled much better for aircraft than for ground-based measurements, because aircraft speeds are typically an order of magnitude larger than the wind advection speed (Lenschow and Stankov, 1986; Adler et al., 2019). Aircraft measurements are time-based and conducted with a true air speed TAS. Employing Taylor's hypothesis (Taylor, 1938), the observed frequency f of flow features in the time series measurement can be related to a spatial wavelength λ (the distance covered by the flow feature) through:

$$f = \frac{TAS}{\lambda}. \tag{2.5}$$

With these concepts in mind, the appropriate averaging length for the calculation of bulk statistical properties for turbulence still needs to be determined. In order to do so, the spectrum of turbulence allows for helpful insight.

The spectra of turbulence

The spectrum of turbulence provides the frequency-resolved variance contribution to the overall variance based on Fourier analysis (its determination is discussed in Sec. 3.4). In the case of atmospheric turbulence spectra, a 4-D spectral tensor is underlying the 1-D (in-situ, 3 components) or 2-D (ADL, 1 component) measurements of wind component spectra (Kristensen et al., 1989). Thereby, the investigated spectrum provides only a limited excerpt of the real turbulence at hand. Nevertheless, spectra reveal important information about the investigated time series and underlying turbulence.

The theoretical assumptions about BL turbulence entail the existence of three distinct regimes, which can be detected from the spectra (Fig. 2.2). For the vertical wind inside the BL, a spectral peak (maximum variance contribution) is expected to exist at low frequencies (long wavelengths, using Taylor's hypothesis for conversion). The peak is associated with the mechanical and convective production of variance by turbulence at or beyond the scale of the BL (hundreds of meters to kilometers). Consequently, the largest scale is called production range and the spectral peak wavelength λ_m is another important parameter in BL research. The energy produced at the largest scale is expected to cascade down towards the smallest scales (millimeters to sub-millimeter), where it is dissipated into heat. Consequently, the smallest scale is called dissipation range. In between production and dissipation range lies the inertial subrange. In the inertial subrange, energy is neither produced nor dissipated, as an equilibrium between production at the largest scale and dissipation at the smallest scale is assumed. Thereby, in the inertial subrange, the turbulent eddies are solely produced through the decay of larger eddies and decay themselves into smaller eddies again until the dissipation range is reached. According to theory, if the process of turbulence decay is fully developed, a characteristic decay of the spectral lev-

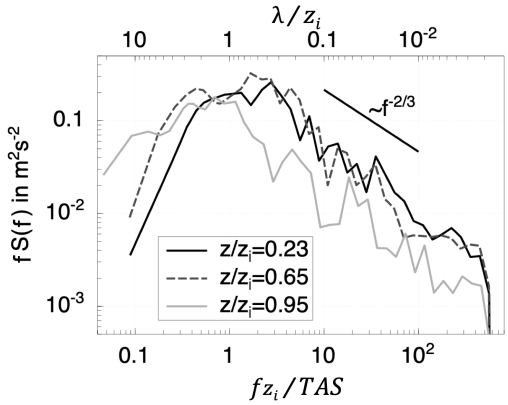

Fig. 2.2.: Aircraft measured energy spectra of the vertical wind variance at three heights in the BL from Hasel (2006).

els towards higher frequencies is expected in the inertial subrange of the spectrum (also called $Kolmogorov's\ law$, a -5/3 slope in the case of $log\ S(f)$ vs. $log\ f$ display, a -2/3 slope in the case of $log\ [fS(f)]$ vs. $log\ f$ display).
An important effect occurs in the spectra for wavelengths larger than the production range. In this region (>O(10 km), >O(1 h)), often a drop-off in the spectral energy level of turbulent fluxes is observable, which is also termed spectral gap (Fiedler and Panofsky, 1970). The spectral gap exists because the next major contributors to energy input into the atmosphere through turbulent fluxes are the diurnal cycle of the atmosphere and synoptic systems, but these provide their input at scales very much larger than the production range. Hence, an intermediate range exists where lower values of energy input by turbulent fluxes can be observed. As outlined and shown by Lenschow and Sun (2007), it must be noted that the spectral gap is present for turbulent fluxes, but not necessarily in the spectra of the wind components (especially the horizontal components) or scalar values. Further, a variety of meso-scale flow phenomena exist which can prevent the drop-off in the spectra (e.g. terrain height variations, gravity waves, longitudinal rolls, cellular convection, dynamical effects induced by clouds and shading of the ground). Nevertheless, the spectral gap provides an approach to separate turbulent flow from mean flow, based on spectral analysis of the energy input provided to the atmosphere through turbulent fluxes.

The choice of an appropriate averaging interval and measurement frequency

Knowledge of the spectra and assumptions used in the determination of bulk statistical properties now allows to identify an appropriate averaging interval and measurement frequency. For reliable variance measurements, the measurement system must enable resolution of all relevant variance contributions from the production range down to the end of the inertial subrange. Thus, a spatial resolution on the order of meters needs to be achieved (requiring a high sampling frequency for fast aircraft measurements). A slightly insufficient measurement resolution (e.g O(10 m)) does not automatically imply useless measurements, as theoretical models describing the behavior of the turbulent spectra for small wavelengths allow for a reconstruction of the lost variance to some extent (Sec. 3.1 and 6.2.VII).
Defining an appropriate averaging interval is slightly more complicated. On the one hand, the averaging interval must be short enough to allow for the valid assumption of homogeneous turbulence and Taylor's hypothesis. On the other hand, the averaging interval must be long enough to capture the contributions of the largest vortices to the spectrum of turbulence reliably. Thus, the averaging interval depends on the turbulence present. It is therefore helpful to

define a measure when neighboring observation become independent. This is achieved through definition of the integral scale $\mathcal{T}$, which is another important measure in BL research as it characterizes the memory of the flow (its calculation is discussed in Sec. 3.4). Typically, an averaging interval which extends beyond the spectral maximum (towards the spectral gap) and covers hundreds of times the integral scale to allow for an adequate statistical basis is seen as appropriate (Lenschow et al., 1994). Even if an appropriate averaging interval is chosen, uncertainties remain because the desired ensemble average has to be approximated by the temporal or spatial average available through measurements (discussed in Sec. 3.4 and the associated Appendix A.9). For example, an appropriate averaging interval typically only allows for sampling of a few of the largest vortices, whose contribution to the overall variance is very large, thus introducing statistical uncertainty.
In summary, this section motivates and describes the concept behind observations of the BL wind field, consisting of two main tasks:

1. Measurement of the mean flow and

2. Characterization of the statistical properties of turbulence.

The first requires reliable identification of the mean flow despite the presence of turbulence. The second requires the same to begin with, but then further the capability for high resolution observations to capture the turbulent deviations in detail across a wide range of scales. Then, the state of the BL turbulence can be characterized through observation of the BL height z_i and the vertical profiles of the vertical wind variance w'^2, the spectral peak wavelength λ_m and the integral scale $\mathcal{T}$. Due to the high accuracy and resolution needed, observations of the turbulent BL wind field are challenging to execute, an overview on existing approaches and systems is provided in the next section.

2.2. Vertically resolved wind observations

Wind observations aspire to capture the current flow state of the atmosphere through measurements, with the following discussion focused on the BL. Associated with the different scales of flow, a variety of observing systems exists which utilize different technologies. While many observations are made at or close to the surface by meteorological ground stations and towers (which are not discussed here for the sake of brevity), a few observation system are able to provide vertically resolved profiles of the wind throughout the BL. Even fewer can provide vertically and spatially resolved observations. Of these, yet even fewer have the ability to resolve turbulent processes. Up to date, no technology exists which is able to capture the full

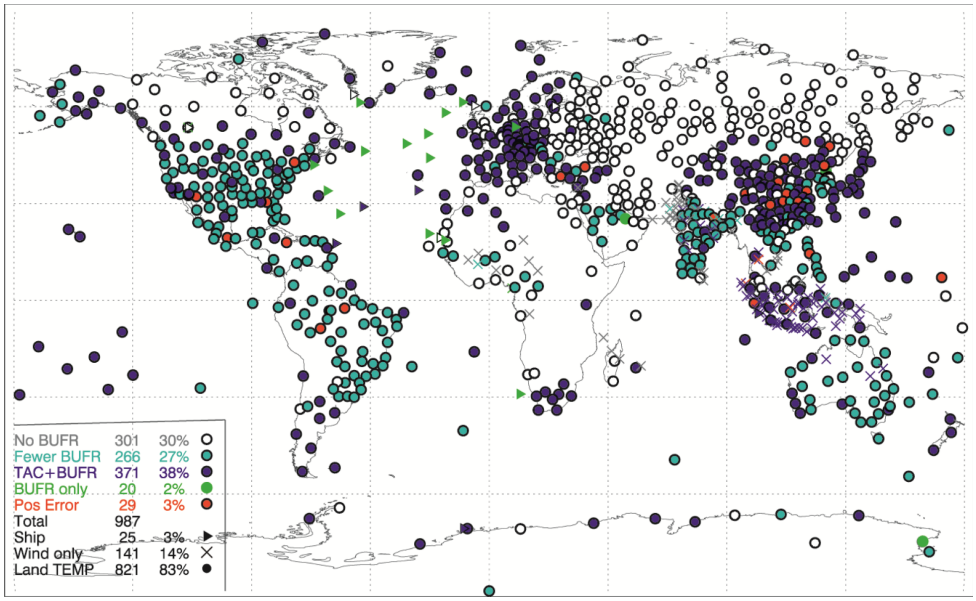

Fig. 2.3.: Global radiosonde observation network from Ingleby et al. (2016). The location of stations for which radiosonde wind reports were processed at the European Centre for Medium-Range Weather Forecasts in the time interval 1–31 Dec 2015 is indicated through circles, triangles and crosses. The color code is non-relevant for the analysis conducted here. © American Meteorological Society. Used with permission.

three-dimensional flow in the atmosphere in turbulent conditions (e.g. three wind components in three dimensions).

Observations of the mean flow

Vertical profiles of the horizontal wind obtained with radiosondes are a backbone of routine operational weather observation for the prediction of the global atmospheric circulation (Ingleby et al., 2016). However, due to the high operational cost the radiosonde network is limited, with almost no observations in remote regions (Fig. 2.3). Additional wind information is provided by aircraft measurements on-board regular passenger flights (Moninger et al., 2003). However, these observations are only conducted along flight routes and vertical information is only available during ascent and descent. Satellite-based AMVs are able to provide wind information also in remote regions and with high temporal resolution (Holmlund, 1998), but their vertical and spatial availability is limited by the required presence of trackable clouds or water vapor features. Radar and sodar based wind profilers offer another way for ground-based remote sensing of atmospheric winds up to the stratosphere. However, they are expensive, e.g. DWD currently only operates four such systems (Deutscher Wetterdienst, 2020b). Operational precipitation radar can also provide winds through radial velocity measurements. These observations require hydrometeors for scattering as well as assumptions about their fall velocity. If sufficient

signal coverage and overlap is available, the combination of multiple radars allows for construction of a three-dimensional mean flow field, as is done over France for example (Bousquet and Tabary, 2014). Sondes flying on tethered balloons are also sometimes used to provide vertical information on the wind, e.g. during research measurement campaigns (Ivey et al., 2014). A recent technology under development is the observation of winds using unmanned aerial vehicles (Barbieri et al., 2019).

Through the last decades, Doppler lidar technology has emerged which enables active remote sensing of atmospheric winds also in clear air (Weitkamp, 2005). Today, the Doppler lidar technology has achieved maturity and the assimilation of Doppler lidar data into operational NWP models is beginning (Kawabata et al., 2014; Sawada et al., 2015). The most common way to infer the horizontal wind from measurements by a single lidar is through scanning in conical circles. The projection of the horizontal wind into the lidar measured radial velocity is then a function of the azimuth angle. Thus, inversion of the measurement knowing the beam pointing-angle allows for retrieval of the horizontal wind (which is hence termed velocity azimuth display, VAD technique, Browning and Wexler, 1968). For research purpose, multiple ground-based lidar systems have been combined in so-called dual- or multi-Doppler retrievals to allow for spatially and vertically resolved observations of the mean flow field (Röhner and Träumner, 2013), also in complex terrain (Fernando et al., 2019; Adler et al., 2020a). Another recent technique is presented by Mauder et al. (2020) in the form of a bistatic lidar. The bistatic lidar consists of one laser source but three spatially separated receiving units, it thus allows for simultaneous measurement of all three wind components in a probe volume by lidar. Recently, the first Doppler lidar (Aeolus) has been launched into space by the European Space Agency after more than 20 years of development (Paffrath et al., 2009; Lux et al., 2018; European Space Agency, 2018). Aeolus now provides the first global wind measurements, which are already assimilated into NWP models and show a positive impact (Deutscher Wetterdienst, 2020c).

An expensive but often used way to obtain vertically resolved wind information for research purpose are aircraft measurements (Lenschow and Stephens, 1980, besides others). A disadvantage is that aircraft measurements only provide wind observations at flight level, if no specific time consuming maneuvers to obtain vertical profiles of the horizontal wind are conducted. For this reason, some aircraft have been equipped with dropsonde launch systems which are also used semi-operationally, e.g. for wind observations in hurricanes (Powell et al., 1998; Dorst, 2007). Unfortunately, launching dropsondes constantly along the flight path is prohibitively expensive and impossible. To allow for continuous spatially resolved wind observations along the flight path nevertheless, some of the used research aircraft have been outfitted with Doppler radar or Doppler lidar systems, which are extensively discussed in the next section.

ADL and airborne Doppler radar (ADR) provide for remote sensing of the wind field with high spatial resolution. ADL retrieved wind profiles have been assimilated into NWP models in multiple experimental studies (Weissmann et al., 2005b; Pu et al., 2010; Weissmann et al., 2012). Similar to ground-based systems, measurement of the horizontal wind is conducted by measuring under multiple azimuth angles (often scanning in conical circles but sometimes also with multiple fixed beams) followed by a retrieval. Hence the name airborne-velocity-azimuth-display, AVAD, is coined, although the retrieval process is better describes as a volume-velocity-processing technique (VVP, see Sec. 3.3). The vertical wind can be explored directly through nadir stares at higher resolution. Nevertheless, some post-processing is required to account for non-exact nadir beam pointing and to retrieve statistical parameters of the vertical wind, e.g. for turbulence measurements.

While up to date published studies on ADL systems are limited to AVAD wind profiling or nadir stare observations, for ADR more sophisticated retrieval strategies such as airborne dual-Doppler have been reported (Leon et al., 2006; Damiani and Haimov, 2006). These techniques enable the retrieval of two wind components in a two-dimensional plane. Two reasons explain the more sophisticated ADR observation strategies compared to ADL: First, ADR systems have been available for much longer and are available on more aircraft, since the radar technology was developed earlier than lidar. Second, changing of the beam direction is more easily done for radar systems, where the whole antenna can be rotated (Hildebrand et al., 1996), multiple antennas can be used simultaneously (Haimov et al., 2018) or the scanning can even be conducted electronically (Lussier et al., 2018). In contrast, for lidar sophisticated mirror-based optical scanner systems are required, which can direct the beam in the desired direction. Mechanical scanning with an optical mirrors adds to cost and complexity as a second system is required and installation outside the aircraft hull (preferential for complex scanning geometries) is difficult due to the needed size of the optical aperture. Up to date, ADL and ADR systems provide the only way to measure the mean horizontal flow with spatial and vertical information in flexible but targeted measurements, as conducted by Weissmann et al. (2005b, 2012).

Turbulence measurements

The presence of turbulence makes wind observations more demanding and challenging. Very few systems are able to provide vertically resolved measurements of turbulence beyond the range of standard flux towers (a few very tall flux towers above a height of 300 m exist, Andreae et al., 2015).

As remote sensing offers a viable and promising opportunity, the development of turbulence retrieval techniques from Doppler lidar and radar measurements has been a focus of research for more than 20 years (Eberhard et al., 1989; Frehlich, 1997; Banakh and Werner, 2005). Today, a number of techniques have reached a level of maturity where reliable results can be obtained regularly. Most commonly applied is the estimation of turbulence characteristics from a single lidar (Newman et al., 2016; Bonin et al., 2017), while other more complex approaches using dual-Doppler retrieval strategies also exist (Röhner and Träumner, 2013). An overview on the available techniques is provided by Sathe and Mann (2013). The bistatic lidar technique mentioned above also allows for the retrieval of turbulent quantities (Mauder et al., 2020). A method for the quantitative measurement of the dissipation rate by aircraft, which can be employed on commercial passenger flights, is presented by Cornman et al. (1995) and Sharman et al. (2014). Dissipation rate measurements are useful for mapping of turbulence levels (e.g. for the warning of other aircraft on high intensity turbulence, especially in clear-air). Still, up to today the most common, reliable and accurate way to obtain vertically resolved information on turbulence are research aircraft. Turbulence measurements by research aircraft have been conducted throughout many decades and provide invaluable information (Lenschow, 1970; Lenschow and Stankov, 1986; Mahrt, 2000).

Along with the outfitting of research aircraft with ADR and ADL systems for mean flow observations, the same systems have been used to retrieve information about turbulence in parallel by a few studies. An overview is provided in the following sections, again it is noticeable that ADR measurements have progressed slightly further than ADL measurements in this respect (Lothon et al., 2005; Strauss et al., 2015). However, turbulence measurements by remote sensing from aircraft provides for a unique set of problems (discussed in Sec. 3.5) and none of the studies so far has addressed all of them simultaneously. For example, a study with correction of the pulse volume averaging (PVA) effect while considering the horizontal wind contamination of the vertical wind measurement, due to non-nadir beam pointing, was missing prior to the work presented here (discussed below).

2.3. Airborne Doppler lidar observations

ADL systems provide unique insight into atmospheric flow phenomena, therefore a variety of systems have been developed throughout the years. An overview of existing systems and experimental validation efforts associated with them is provided here, followed by a few highlight studies and the presentation of a new ADL system concept.

Because ADL presents a combination of aircraft, scanner and lidar systems the data analysis requires adequate processing schemes. The motion correction necessary for airborne wind measurements is generally challenging because the atmospheric contribution to the measured velocities is very small compared to the aircraft contribution. Hence, precise and reliable motion correction of the aircraft contribution is required and small offsets in any of the components can introduce large error. The analysis is complicated further for wind profiling because the lidar is only able to measure the contribution of the atmospheric velocity in beam direction (radial or line-of-sight velocity, LOS). As a result, a retrieval process is required to obtain vertical profiles of the horizontal wind, which can again introduce error. Thus, validation of the achieved measurement accuracy, both with respect to the raw motion corrected radial velocities as well as the retrieved wind profiles, is important.

Validation strategies

Associated with the radial velocity and AVAD wind profile accuracy assessment two different validation approaches exist.

First, the radial velocity accuracy (also termed LOS accuracy) is typically assessed by using the measured ground return velocity (Baidar et al., 2018, are the exception to the rule, as they use simulated radial velocities from dropsonde measurements). In this method, the ground serves as a stationary reference target, hence any ground velocity remaining after motion correction is attributable to error in the motion correction procedure (which is explained in Sec. 3.2). Owing to the direct approach a robust handle on the radial velocity error due to motion correction can be obtained. The ground return even allows for motion correction in case of unavailable or unreliable inertial navigation system (INS) data, as is done by Augere et al. (2019) (if the lidar is able to sense the ground, e.g. no blocking by clouds).

Second, the determination of the wind profiling accuracy is more complicated due to the retrieval process involved. A common method to assess the reliability of retrieved Doppler lidar wind profiles is by comparison to wind profiles obtained using other measurement systems. However, a number of problems exist for these 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. Systems with similar remote sensing characteristics, e.g. ADRs, also suffer from co-location problems and are prone to the same retrieval errors as ADL. Consequently, results can show large differences (De Wekker et al., 2012; Tian et al., 2015; Bucci et al., 2018). Therefore, the

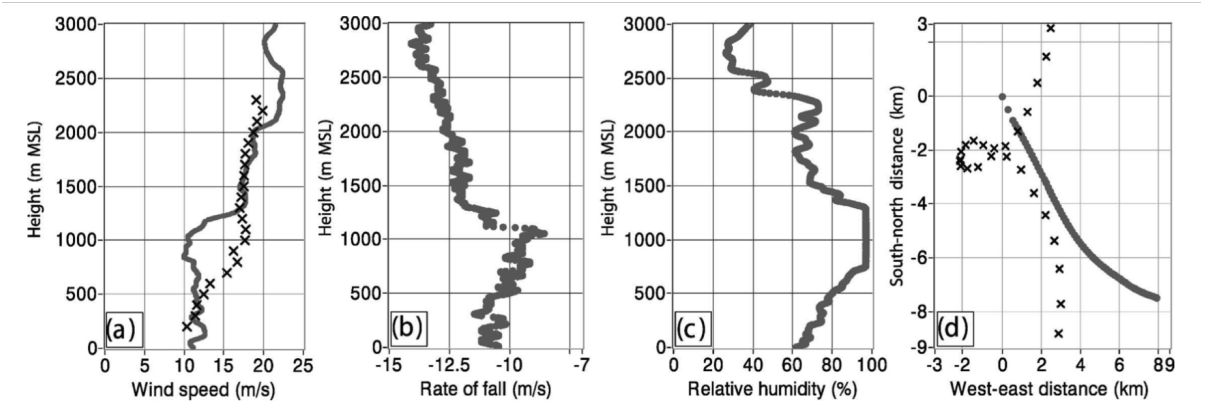

Fig. 2.4.: Example comparison of ADL retrieved wind profile (crosses) with dropsonde measurements (points) from Weissmann et al. (2005b). Shown are: (a) Wind speed, (b) rate of fall, (c) relative humidity and (d) horizontal positions of dropsonde and lidar radial velocity measurements used for retrieval. © American Meteorological Society. Used with permission.

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 (Reitebuch et al., 2001; Weissmann et al., 2005b; Chouza et al., 2016a; Bucci et al., 2018; Greco et al., 2020). Nevertheless, both systems still exhibit very different sampling characteristics and volumes. This difference often makes a direct comparison of the results challenging, as it is difficult to determine if the observed deviations occur due to the differing sampling volumes, violation of the flow homogeneity assumption used for retrieval or actual instrument error. An example from Weissmann et al. (2005b) is provided in Fig. 2.4. Despite the well-matching location between ADL retrieved wind profile and dropsonde measurement, differences greater than $5\,\mathrm{m\,s^{-1}}$ are present in the wind profiles below 1300 m. Weissmann et al. (2005b) attribute these to a fall of the dropsonde through a convective cloud updraft based on the reduced fall speed and 100 % relative humidity. Thereby, the different sampling characteristics and resulting spatio-temporal co-location error prevent a reliable identification of instrument or retrieval errors. This problem is amplified in the BL, where turbulent flow leads to a permanent violation of the flow homogeneity assumption used for AVAD retrieval and large differences already for small shifts in location.

Due to the challenges associated with ADL measurements, theoretical analysis of the measurement and retrieval procedure (Sec. 2.4) presents another important possibility for measurement accuracy evaluation besides the experimental validation discussed here.

2.3.1. Overview of active systems and experimental validation efforts

The following discussion includes ADLs on which studies were published in the years 2010-2020. While older systems do exist they are not operational anymore, an overview of them can be found in Weissmann (2006). Fig. 2.5 provides an overview of the typical measurement operation envelope of the aircraft associated with the ADL systems. Further, the measurement strategies for which results have been published are also shown, along with an overview of quantitative evaluation available in literature. Details on the systems can be found in Tab. 2.1 for the system hardware and Tab. 2.2 for the associated validation efforts. Validation as used here refers to the publishing of quantitative information on the achieved system accuracy (see also Appendix A.1).

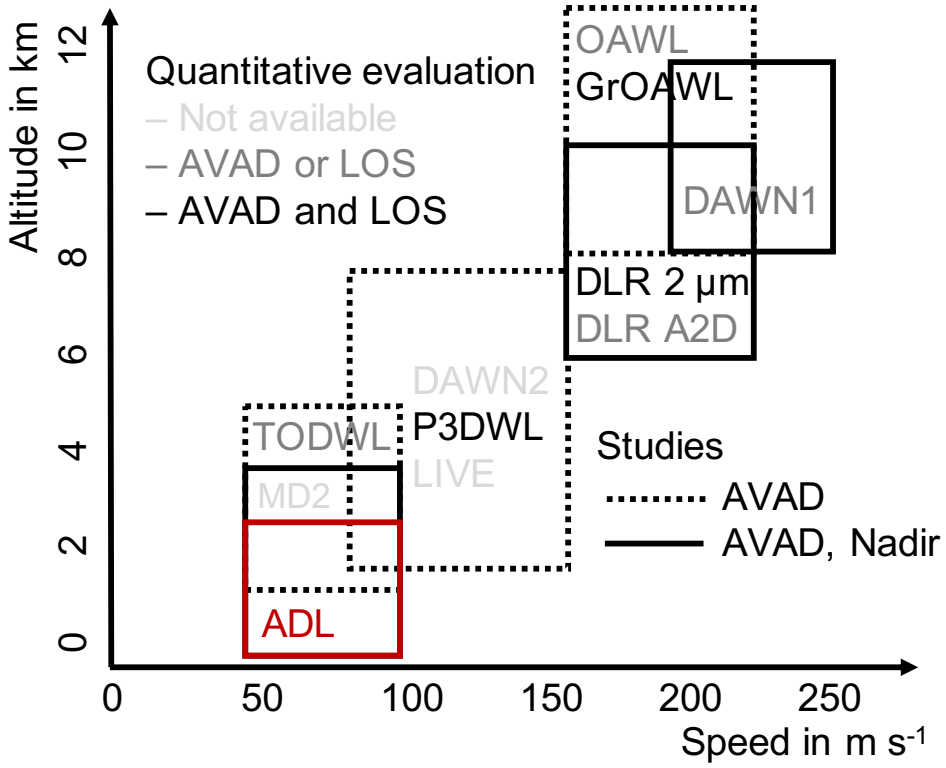

Fig. 2.5.: Overview of ADL systems for which studies were published in the years 2010-2020. The boxes indicate the typical measurement operation envelopes of the research aircraft when flying the respective ADL systems. The names refer to the acronyms given in Tab. 2.1. The line style shows for which scanning capabilities studies are documented in peer-reviewed literature. The color code shows for which measurements a quantitative evaluation is documented in peer-reviewed literature. The red box outlines the operation envelope of the Dornier 128-6 (call-sign D-IBUF) research aircraft with the original system intended for usage in the work presented here and simulated in the ADLS.

The by far most utilized ADL system is a 2 μm Doppler lidar (DLR 2 μm) based on a Falcon 20 aircraft operated by the German Aerospace Center (DLR, Deutsches Zentrum für Luft- und

Raumfahrt e.V.) (Witschas et al., 2017, and references therein). The Falcon 20 has a typical aircraft speed of 200 m s^{-1} during measurement flight at a flight altitude between 6 – 10 km. It is routinely operated under instrument flight rules (IFR). The lidar operates at a wavelength of 2022 nm with a pulse repetition frequency (PRF) of 500 Hz, the utilized data output rate is 1 Hz. The DLR 2 μm scans with a typical scan rotation duration of 30 s (stop-and-stare is possible) with a scan elevation angle of 70° (20° off-nadir). Thereby, vertical profiles of the horizontal wind can be obtained every 6 km along the flight path. The DLR 2 μm is also able to measure the vertical wind through nadir stares, for which multiple studies have been reported. Chouza et al. (2016b) assess the measured ground return velocities (radial velocity error) quantitatively and conduct a beam pointing-angle calibration procedure based on it. They report a slight dependence of the beam pointing-angle on cabin pressure. Further, Chouza et al. (2016b) and Witschas et al. (2017) analyze the accuracy of the nadir measurements based on theoretical considerations, the measured ground return velocities and through the difference between up- and downwind flights. Multiple accuracy assessments of the retrieved vertical profiles of the horizontal wind have been conducted by Weissmann et al. (2005b), Chouza et al. (2016a), Reitebuch et al. (2017) and Schäfler et al. (2020) in comparison to dropsondes. A noticeable difference in the reported accuracy exists between the studies (Tab. 2.2). The differences reported are too large to be explained by varying measurement system accuracy, as the radial velocity accuracy is expected to be similar between the studies (the used system components are the same and the ground return is routinely checked by the other studies as well, showing no gross motion correction errors). Thus, the difference emphasizes the difficulty of conducting wind profiling accuracy evaluation and the impact of factors beyond the pure system accuracy, such as the retrieval strategy and the flow present in the retrieval volume.

The Falcon 20 can host another ADL in parallel to the DLR 2 μm, which is a direction detection lidar (DLR A2D). The DLR A2D serves as an airborne demonstration and calibration system for the Aeolus satellite mission (Lux et al., 2020, and references therein). The lidar operates at a wavelength of 354 nm with a PRF of 50 Hz, the utilized data output rate is 1/18 Hz. Differing from the usual coherent detection signal processing, the direct detection uses interferometers to determine the Doppler frequency shift between outgoing and returning signal. It thereby allows for processing of both aerosol (Rayleigh channel, double-edge technique using two bandpass filters in two Fabry–Pérot interferometers) and cloud (Mie channel, using a fringe-imaging technique in a Fizeau interferometer) return signals simultaneously. However, an accurate ground return calibration (termed zero-wind correction in this case) is needed for A2D radial velocity measurements (Lux et al., 2018). The lidar measures in a stare mode at an off-nadir angle

of 20°, hence no full AVAD or nadir measurements are reported. The DLR 2 μm is used as a reference system for accuracy assessment of the A2D.

Another satellite demonstrator mission (termed OAWL) with a fixed-beam laser is introduced by Tucker et al. (2018) onboard the high-flying NASA WB-57 aircraft. The OAWL system uses a laser source at 355 nm with an optical auto-covariance detection technique for signal analysis. The OAWL radial velocity accuracy is controlled using the measured ground return velocities (Tucker et al., 2018), but wind profiling accuracy is only briefly validated against a few measurements from a radar wind profiler (as it requires turning of the aircraft). A further development stage on the same aircraft led to the GrOAWL system, which utilizes two lasers at a wavelength of 532 nm and 90° azimuth separation, allowing for AVAD retrieval. Baidar et al. (2018) validate both the radial velocity accuracy and the retrieved wind profiles against 21 dropsondes (whose measurements are used to simulate radial velocities).

Another ADL based on a high and fast flying aircraft is the DAWN1 lidar onboard the NASA DC-8 (Greco et al., 2020, and references therein). The DAWN lidar emits more powerful individual shots (leading to the damage of an optical mirror through burning of settled dust, as reported in Kavaya et al., 2014) but at a lower PRF, the available data rate is comparable to other systems with 0.5 Hz. So far, the motion correction accuracy has been provided only based on the ground return measurements during 5 AVAD scans for the DAWN1. For this system AVAD horizontal wind profiling has been reported with 2- (separated by 90°) or 5-point (separated by 22.5°) conical stare measurements at 30° off-nadir angle. An extensive accuracy assessment of retrieved wind profiles in comparison to dropsonde data is conducted in Greco et al. (2020). Nadir measurements geometries have also been reported by Turk et al. (2020), however without providing insight into the achieved measurement accuracy.

The same lidar system has also been mounted and operated on a smaller and slower UC-12B aircraft (DAWN2), which is a modified Beechcraft B200 King Air aircraft, for two case studies (Koch et al., 2014; Beyon et al., 2014, besides more grey literature referenced therein). With unchanged lidar characteristics, the UC-12B DAWN2 operates at a typical aircraft speed of 135 $\mathrm{m\,s^{-1}}$ and a flight altitude of 1.5−6 km. DAWN2 also uses an AVAD scan pattern limited to two points stares, separated by 90° azimuth at 30° off-nadir angle. Koch et al. (2014) state that a radial velocity accuracy assessment is conducted by analysis of measured ground return signals, but without providing further details except that it is better than 1 $\mathrm{m\,s^{-1}}$. Two AVAD retrieved wind profiles are compared to those from a radar wind profiler and a dropsonde, also without further accuracy quantification. Results on nadir measurements have not been published for the DAWN2.

Tab. 2.1.: Overview of existing ADL systems on which studies were published in the years 2010-2020.

System	Agency	Aircraft	Altitude in km	Speed in $m\,s^{-1}$	Wavelength in nm	Pulse energy in mJ	PRF in Hz	Data rate in Hz	Pulse length in ns	Scans strategies published
DLR 2 μm	DLR	Falcon 20	8-12	200	2022	1-2	500	1	400	AVAD, nadir
A2D	DLR	Falcon 20	8-12	200	354	60	50	1/18	20	Stare
OAWL	NASA/Ball	WB-57	8.5-12 (up to 20)	200	355	20	200	1/10	18	Stare, AVAD
GrOAWL	NASA/Ball	WB-57	8.5-12 (up to 20)	200	532	1.5	200	1/10	≈25	AVAD
DAWN1	NASA	DC 8	8-12	250	2053	100	10	0.5	180	AVAD, nadir
DAWN2	NASA	UC-12B	1.5-8	135	2053	100	10	0.5	180	AVAD
P3DWL	NOAA, AOML	WP-P3D	2-7	130	1617	1.5	500	1 (up to 10)	320	AVAD
LIVE	ONERA/ SAFIRE	ATR-42	2-7	100	1500	0.3	14000	3	700	AVAD
TODWL	US Navy/ UVa+SWA	Twin Otter	0.5-3	60	2022	2	100	1	320	AVAD
MD2	NOAA, ESRL	Twin Otter	0.5-3	60	1543	several 10^{-3}	10000-20000	< 10	400 (166-1000)	AVAD, nadir

Tab. 2.2.: Overview of LOS and AVAD accuracy validation studies associated with the systems in Tab. 2.1.

System	LOS validation	Method	Error in $m\,s^{-1}$		AVAD validation	Method	Points (Profiles)	Error in $m\,s^{-1}$	
			Syst.	Rand.				Syst.	Rand.
DLR 2 µm	Chouza et al. (2016b)	All ground returns	0.036	0.1	Weissmann et al. (2005b) Chouza et al. (2016a) Reitebuch et al. (2017) Schäfler et al. (2020)	Dropsonde	740 (33) 1329 (22) 938 (15) 529 (15)	< 0.01 0.08 -0.03 0.05	1.20 0.92 1.46 1.87
A2D	Lux et al. (2018)	DLR 2 µm LOS	-0.21	2.0	-	-	-	-	-
OAWL	Tucker et al. (2018)	All ground returns	0.04	3.4	Tucker et al. (2018)	Radar profiler	≈ 600 (6)	n/a	n/a
GrOAWL	Baidar et al. (2018)	Dropsonde LOS	0.07	1.04	Baidar et al. (2018)	Dropsonde	≈ 521 (21)	0.28	1.38
DAWN1	Greco et al. (2020)	5 scan ground return	< 0.25	≈1	Greco et al. (2020)	Dropsonde	16207 (162)	0.19	1.61
DAWN2	Koch et al. (2014)	All ground returns	n/a	< 1	Beyon et al. (2014)	Dropsonde, radar profiler	≈ 100 (2)	n/a	n/a
P3DWL	Bucci et al. (2018)	n/a	n/a	n/a	Bucci et al. (2018)	Dropsonde in-situ aircraft ADR	2056 (49) 469 40559	0.01 0.55 -0.04	2.37 3.31 4.13
LIVE	Augere et al. (2019)	All ground returns	n/a	n/a	Augere et al. (2019)	Ground lidar	≈ 30 (1)	n/a	n/a
TODWL	Godwin et al. (2012)	100 ground returns	n/a	< 0.2	De Wekker et al. (2012)	Radar profiler	≈ 200 (≈ 10)	n/a	< 0.5
MD2	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

A system with similar operating characteristics as the DAWN2 is the P3DWL onboard the NOAA's Lockheed WP-3D Orion presented in Pu et al. (2010); Zhang et al. (2018) and Bucci et al. (2018). The P3DWL has been used to obtain wind profiles in and around tropical hurricanes, nadir measurements have not been reported. The system uses the same lidar as in this work, a 1.6 μm Lockheed Martin Coherent Technologies WTX. For AVAD scanning a 2-axis scanner is available which performs a 12 point stop-and-stare pattern at 20° off-nadir angle. The scanner is also able to point upwards. An extensive AVAD performance analysis is conducted by Bucci et al. (2018), who compare the AVAD retrieved wind profiles to 49 dropsondes as well as in-situ measured, microwave radiometer and ADR data from the same aircraft. No results on the achieved radial velocity accuracy have been published.

Recently, another ADL system termed LIVE has been introduced by Augere et al. (2019), based on the French SAFIRE ATR-42 aircraft with similar aircraft characteristics as the WP-3D. Noteworthy is the LIVE laser unit, which is an all-fibre laser at a wavelength of 1.5 μm and the associated high PRF of 14 kHz. Scanning is conducted with a conical scan of 17 s duration at 30° off-nadir angle. Augere et al. (2019) report that their aircraft-based INS is not accurate enough to allow for motion correction, hence the measured ground return velocities are used instead. Further insight into the achieved radial velocity accuracy is not provided. AVAD accuracy analysis is conducted through comparison to one wind profile from a ground-based Doppler lidar, also without quantitative evaluation.

The lowest and slowest flying ADL systems up to date are both implemented onboard Twin Otter aircraft. The first such system (TODWL) was presented by De Wekker et al. (2012) and Godwin et al. (2012) for a wind profiling case study in the Salinas valley in California. The TODWL has also been used by Fernando et al. (2015) and Pal et al. (2016), however focused on signal-to-noise ratio (SNR, as a proxy for aerosol content) instead of wind. De Wekker et al. (2012) specify a radial velocity accuracy of $< 0.2\,\mathrm{m\,s^{-1}}$ based on the measurement of 100 ground return velocities, which they also use to conduct a beam-pointing angle calibration. AVAD wind profiling is conducted by a 12-point step-stare scan with 20 s duration at 20° off-nadir angle. Wind profiling accuracy evaluation is conducted through comparison to approximately 10 wind profiles from a nearby ground-based radar wind profiler. They specify an accuracy of $< 0.5\,\mathrm{m\,s^{-1}}$, however without providing further insight into how this value is obtained. Based on the results obtained in the work presented here, their accuracy appears unreasonably high considering that their AVAD wind profiling is conducted in the turbulent BL, where flow inhomogeneity induces retrieval error.

A new system which is currently undergoing system tests is the NOAA Earth System Research Laboratory (ESRL) developed MD2, also onboard a Twin Otter (Schroeder et al., 2020). Note-

worthy is its rather compact and lightweight design compared to others, further it exhibits a very high PRF similar to LIVE. Both AVAD (up to $60\,^{\circ}\,s^{-1}$ scan speed) and nadir scanning geometries are possible with high lidar measurement frequency (up to 10 Hz). An accuracy assessment is not conducted by Schroeder et al. (2020), as the focus is on system introduction and capability presentation. Consequently, no insight into the achievable measurement accuracy for BL measurements is available yet.

2.3.2. Application studies

ADL systems have been used to study a variety of atmospheric phenomena around the globe, among them gravity waves (Chouza et al., 2016b; Witschas et al., 2017), hurricanes (Bucci et al., 2018; Zhang et al., 2018), mid-latitude storms (Weissmann et al., 2005b; Schäfler et al., 2020), aerosol transport (Chouza et al., 2016a), convective systems (Turk et al., 2020), aircraft wake vortex (Rahm et al., 2007), BL turbulent fluxes (Kiemle et al., 2007, 2011) as well as meso-scale flow features across a valley (De Wekker et al., 2012), a wind farm (Koch et al., 2014) and a wild fire (Schroeder et al., 2020). For the sake of brevity, only recent studies with an investigation component linked to the BL or of importance to the work presented here are discussed briefly in the following. For an overview of older studies (before the year 2005) the reader is referred to Weissmann (2006).

The first study presenting ADL measurements of BL flow is by Kiemle et al. (2007), who use a pre-cursor to the DLR 2 μm system (similar in characteristics). The Doppler lidar is co-located with a differential absorption lidar (DIAL) humidity profiling lidar in their study, with both system staring nadir. Despite the rather coarse resolution of 1 s (corresponding to 200 m) for both lidars, the joint dataset is used for the calculation of latent heat flux profiles in the BL using the eddy-covariance method. While they calculate sampling uncertainties due to the finite time series, instrument noise and co-location offset between DIAL and Doppler lidar, other effects such as motion correction accuracy (for which they use a refined motion correction procedure using the measured ground return velocities), PVA and horizontal wind contamination due to non-nadir beam pointing are not evaluated quantitatively. Especially the neglect of PVA appears problematic given the coarse resolution of both systems. Nevertheless, Kiemle et al. (2007) are able to document important BL features such as a BL drying due to a net upward humidity transport and entrainment of dry air from above. Together with a water vapor budget analysis they report that the drying inhibited cloud growth and prevented deep convection.

A similar study is repeated for the BL over complex terrain (the black forest) by Kiemle et al. (2011) using the DLR 2 μm system. They apply a similar analysis scheme as Kiemle et al.

(2007) but with some refinements and more detailed investigations of the statistical uncertainties. Further, they give detailed consideration to inevitable lidar data gaps and their effect on the calculation of variances and spectra. An evaluation of the ADL measured fluxes and spectra is conducted through comparison to in-situ measured values by the Dornier 128-6 aircraft also used for this work, with good agreement reported within the uncertainty bounds calculated by them. Hence, Kiemle et al. (2011) are able to provide vertically resolved latent heat flux measurements above complex terrain, a previously inaccessible quantity. They report significant variation between flight legs but almost constant flux profiles vertically. Besides others, the Kiemle et al. (2007) and Kiemle et al. (2011) studies provide an important foundation for the analysis of the real-world ADL data conducted as a part of the work presented here.

A first study of meso-scale flow phenomena using a slow-flying aircraft is conducted by De Wekker et al. (2012) using the TODWL. De Wekker et al. (2012) conduct AVAD wind profiling with an along-track resolution of 1.5 km to investigate the spatio-temporal structure of BL flow across the Salinas valley in California. Thus, they are able to measure the advance of an accelerating sea-breeze front turning into an upvalley flow. Detailed insight into local flow phenomena such as jet-like structures and cross-valley winds is provided. The companion study by Godwin et al. (2012) utilizes the same ADL data. Their study proposes a method to extend the measurement of ADL velocities towards the ground by an additional 100 m. This extension is achieved through tracking the aerosol spectral peak in the Doppler signal, enabling the removal of the high intensity ground return signal.

Although conducted by ADR (which requires the presence of clouds or rain in the BL, not the usual case), the study of Strauss et al. (2015) also warrants mentioning here, as it presents a number of concepts relevant for this work. Building upon a study by Geerts et al. (2011), Strauss et al. (2015) analyze high resolution ADR measurements of the vertical wind in the BL in the lee of a mountain range. Thereby, they are able to identify flow phenomena such as breaking mountain waves and atmospheric rotors. Further, they provide an extensive uncertainty estimation framework for quantitative vertical wind variance retrievals using a fixed-beam system. This includes an estimation of the horizontal wind contamination effect due to non-nadir beam pointing, which is also used in this work (Sec. 3.5). However, while they take into consideration the effect of PVA on the variance retrieval, they do not correct its systematic effect or quantify it. Further, they do not take into account the effect of systematic and random uncertainties due to finite measurement lengths, as outlined by Lenschow and Stankov (1986).

Another study of relevance to this work is presented by Chouza et al. (2016b), who investigate gravity waves induced by the flow over an island using the DLR 2 μm system. They remove the horizontal wind contribution from the nadir measured radial velocity based on wind profiles

obtained from wind profiling in the vicinity. Thereby, they are able to distill the vertical wind component despite a non-nadir pointing beam. Their procedure is also used by Witschas et al. (2017), but based on model data. Both studies are able to provide detailed insight into the characteristics of obstacle-induced gravity waves through the analysis of wavelength, -amplitude and -propagation as well as spectral decomposition. The horizontal wind removal as conducted by Chouza et al. (2016b) is also used in the work presented here, to reduce the horizontal wind contamination effect on the measured vertical wind due to non-nadir beam pointing.
Recently, the new M2D ADL was presented by Schroeder et al. (2020), which is thematically and design-wise the closest to the ADL used in this work. Their study gives detailed insight into the technical specifications. Both an AVAD and nadir application example for wind measurements in the vicinity of wild fires are presented. However, an in-depth analysis of the achieved measurement accuracy and their potential for meteorological evaluation is not conducted so far, as their application analysis is based on quicklooks.
In summary, three things become clear from the conducted system and research overview:

1. There is a need for further enhancement of ADL capabilities with respect to BL measurements through the development of high-resolution lidar onboard slow-flying aircraft.

2. Measurement system accuracy validation is an important part of ADL systems, although the amount and quality of validation conducted strongly varies between studies. Further studies and new approaches to evaluate ADL measurement accuracy and potential for BL measurements are needed, both for wind profiling and nadir measurement approaches.

3. A quantitative estimation of the nadir vertical wind variance measurement accuracy, considering all sources of error and estimating their associated uncertainties (detailed in Sec. 3.5), has not been conducted for ADL measurements up to date.

2.3.3. A new airborne Doppler lidar for boundary layer research

The work presented here is dedicated to the development of a new scanning ADL for BL research. Although already 10 years in the making, serious project development only began with this thesis. Unfortunately, the scanner could not be readied in-time for the measurements conducted within this work, due to unforeseen development issues (e.g. weight and rigidity problems of the scanner boom, prolonged certification requiring multiple test flights) encountered by the scanner manufacturing company. Hence, real-world measurements obtained using a non-scanning, fixed-beam nadir staring prototype ADL version are analyzed in Chap. 6. For the prototype, scanning for wind profile retrieval is conducted by banking the aircraft in turns.

Thereby, fundamental insight into high-resolution ADL BL measurements can be gained and the measurement quality evaluation procedures developed are directly transferable to a scanning system.

The full system is available for measurements with completion of this thesis. Hence, and due to its similarity to other available ADL systems, a full and flexible scanning system is investigated in the theoretical section of this work using the ADLS. A short introduction of the full ADL including the scanner is presented here, because it forms the basis of the theoretical section in this work (while only the aircraft and lidar are combined for the prototype ADL).

Aircraft

The Dornier 128-6 (Do128, call sign D-IBUF) is a medium-range turboprop research aircraft operated by the Institute of Flight Guidance at the Technische Universität Braunschweig, Germany (Fig. 2.6). The Do128 is a high-wing, fixed-gear aircraft with two propeller turbines and short take-off and landing capabilities. It operates at a typical measurement speed of 65 $\mathrm{m\,s^{-1}}$ with a flight duration of up to four hours. While it is certified for instrument flight conditions it is operated under VFR for the ADL installation with a limited ceiling of up to 3000 m. Towards the surface, the Do128 is certified to fly as low as 100 m above land (Sec. 6.2.VII) and 30 m above sea. The usual crew consists of two pilots, a flight engineer and a flight scientist. The VFR operation enables an adjustment of flight plans in flight, which allows for maximum flexibility and targeted measurements. The Do128 hosts extensive measuring equipment for meteorological experiments, including a 100 Hz turbulence probing system for wind vector, temperature and humidity measurements (Corsmeier et al., 2001). The turbulence probe enables the calculation of turbulent fluxes, which has been frequently utilized by previous studies (Grunwald et al., 1996; Hasel et al., 2005; Kiemle et al., 2011; Platis et al., 2018).

Fig. 2.6.: The Dornier 128-6 (D-IBUF) in flight. The white nose-boom protruding from the aircraft is visible.

The three-dimensional wind vector is measured through pressure transducers in a 5-hole probe. The 5-hole probe is located at a nose-boom in the undisturbed airflow in front of the aircraft (Fig. 2.6). The measured wind velocity has to be corrected for the aircraft velocity to obtain

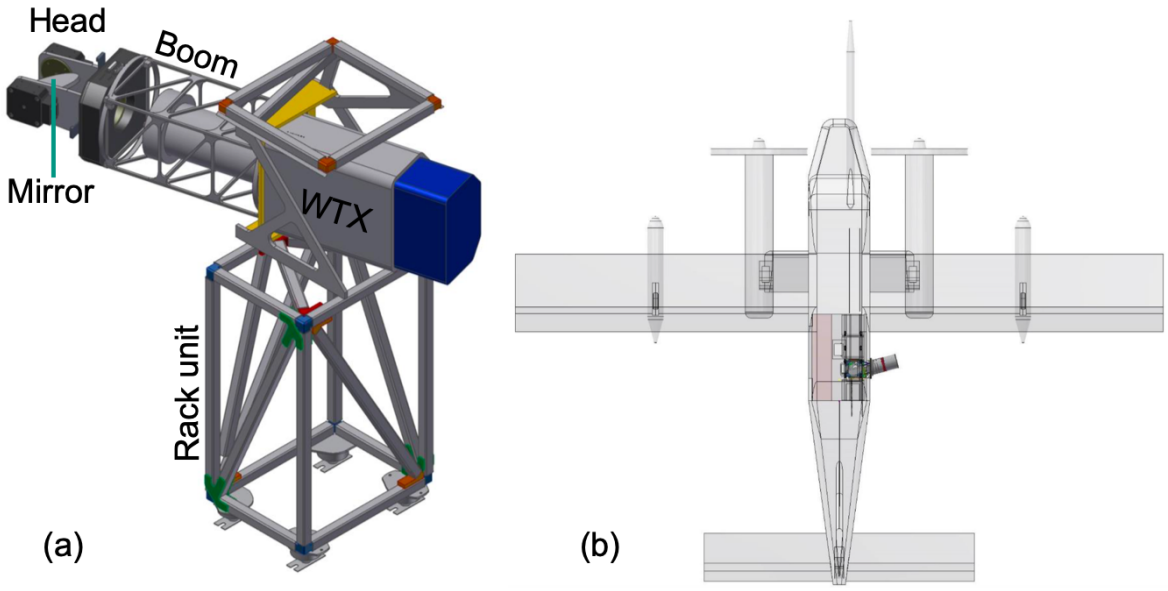

Fig. 2.7.: Arges scanner rack unit (a) and its location inside the aircraft (b). The WTX transceiver is mounted to the scanner rack unit which extends outside the aircraft fuselage. The scanner head is mounted at the end of the scanner boom and includes the mirror which directs the laser beam in the atmosphere.

the wind in an earth-related coordinate system. This is done with the help of an INS, which provides the acceleration, velocity and position of the aircraft with more than 100 Hz data rate. Since this gyroscopic measurement instrument can be subject to drift it is stabilized with the data from a GPS navigation receiver.

The INS data is also used for the raw motion correction of the acquired lidar data. However, analysis of the lidar measured ground return velocities reveals non-negligible errors in the provided INS data. Hence, the lidar data is subjected to a further refined motion correction procedure (Sec. 6.2.IV).

Scanner

The 2-axis scanner system is designed and manufactured by the Bavarian company Arges (www.arges.de). The scanner is situated in its own dedicated rack unit in the rear of the aircraft (Fig. 2.7). A scanner boom of 0.8 m length extends from inside the cabin through a window to the outside and is covered by an aluminum sheeting. Situated at the end of the boom is the scanner head, which allows for 350° elevation movement (the blanked 10° sector is towards the wing). Hence, the scanner allows for measurements of both hemispheres, thereby also permitting flight close to the ground with an upward pointing lidar. Located inside the scanner head

is a 10 cm aperture mirror, optimized for reflection at 1617 nm. The mirror allows for ±30° movement in azimuth direction at all elevation angles. The mirror is covered by a protective glass attached to the scanner hull, which is also optimized for transmission at 1617 nm.
The unit is driven by the Arges scanner controller, which sends encoded signals to the scan motors for movement and allows for user interaction. The maximum scan speed is $30\,^{\circ}\,s^{-1}$. The scan encoders record azimuth and elevation angles which are stored in the controller. The controller is connected to the aircraft data system and receives the current aircraft orientation (pitch, roll, yaw angles). Thus, a scan mode which corrects for aircraft movement with 20 Hz update rate is available, which is valuable as it allows for constant nadir measurements.
While a basic functionality and certification of the scanner system has been achieved recently, a full investigation on the achieved system performance and pointing accuracy is still to be conducted pending the first measurement flights.

Lidar

The lidar used in this work is the newest generation of transceivers (WTX) produced by Lockheed Martin Coherent Technologies. Lidar systems with similar characteristics are often used in ADL studies (Weissmann et al., 2005b; De Wekker et al., 2012; Chouza et al., 2015; Witschas et al., 2017). An equal system has been used by Bucci et al. (2018) and Zhang et al. (2018). The WTX operates at a wavelength of 1617 nm, which provides for class 1M eye-safe measurements at the given beam characteristics. The WTX has a Gaussian pulse width of $\sigma_p = 1.8 \times 10^{-7}$ s, the pulse energy is 2.7 mJ with a PRF of 750 Hz. To reduce random noise in the radial velocity estimation the system averages between $75 - 750$ pulse. Accordingly, the data output rate can be varied in the range $1 - 10$ Hz. For maximum resolution a 10 Hz data output rate is used in this work, which is shown to produce reliable radial velocity measurements with little uncorrelated noise in Sec. 6.2.III. Further, a range gate length of $\Delta p = 69.53$ m with a non-overlapping spacing of $\Delta R_0 = 69.53$ m (the distance between subsequent range gate centers) is used for data processing, starting at a distance of 400 m from the lidar. Further information on the lidar and the associated signal processing procedure is provided in Sec. 3.1, a very similar system is described in Grund et al. (2001).
Originally built and used for ground-based studies (Röhner and Träumner, 2013; Träumner et al., 2015), the equipment associated with the lidar has been modified to allow for airborne usage for the work presented here (an overview on the system components is provided in Fig. 6.1). This includes a replacement of the system control computer (SCC) with an airworthy counterpart, upgrading all other components to an airworthy status (flame retardant cable hulls,

securing of internal components against vibration) and modification of the transceiver mounting plate.
The lidar and aircraft do not share a common time server, therefore time-synchronization has to be conducted in post-processing (Sec. 6.2.I). Further, at the moment no aircraft or scanner data is fed into the lidar. Hence, as no frequency tuning of the lidar processing unit is conducted (e.g. through the analog front end, as in Bucci et al., 2018), the maximum folding velocity of the lidar at $38\,\mathrm{m\,s^{-1}}$ has to be considered and limits the allowed beam pointing directions of the scanner.

2.4. Theoretical investigations on wind profiling accuracy

The new system is designed for BL research and allows for wind profiling and nadir investigation of the turbulent BL wind field at high resolution. However, for AVAD wind profiling in the turbulent BL flow a previously insufficiently addressed challenge exists: The violation of the flow homogeneity assumption used for wind profile retrieval is expected to be of non-negligible but unknown magnitude. Due to the challenges associated with evaluating ADL wind profiling accuracy from experimental studies (Sec. 2.3.1 and Tab. 2.2), idealized simulations of Doppler measurement systems are useful. Theoretical study can provide detailed insight into the capabilities and limitations of ADL systems and wind profiling retrieval strategy.
The work presented here introduces a new approach for the investigation of wind profiling accuracy in turbulent flow by means of theoretical study: an LES-based airborne Doppler lidar simulator. In order to put the ADLS into perspective this section provides a brief overview of previously existing theoretical studies on wind profiling accuracy. Because the wind profiling problems encountered by ADR as well as ground-based lidar and radar are very similar to those of ADL, studies from all these realms are included in the discussion.
Early studies determined the representativeness of Doppler radar and 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; Koscielny, 1984; 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 (Muschinski et al., 1999; Scipión et al., 2009; Scipion, 2011; Stawiarski et al., 2013; 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 challenges have been investigated with statistical models by Baker et al. (1995); Gamache et al. (1995) and Frehlich (2001). For airborne systems, Reitebuch et al. (2001); Leike et al. (2001); 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 deficiencies. While LES-based studies of airborne measurements for the investigation of errors and uncertainty estimation methods associated with turbulent fluxes exist for 20 years already (Schröter et al., 2000; Sühring and Raasch, 2013; Sühring et al., 2019), an LES-based study of an airborne Doppler system was missing previous to the work presented here. A brief overview of developments relevant for the LES-based investigation of ADL is provided in the following.

2.4.1. Idealized turbulence model based analysis

A first extensive study on the effect of flow inhomogeneity on the retrieved wind profiles is conducted by Koscielny (1984) for ground-based systems. He analytically solves the wind profiling problem considering a linearly changing wind field and three simple scanning geometries (fixed beam, azimuthal scanning and elevation scanning). Koscielny (1984) shows that deviations from the assumed wind field can be separated into two contributions. The first contribution (termed meteorological variance by him) is caused by deviations about the uniform wind on scales larger than the resolution volume due to turbulence. The second contribution is caused by receiver noise and turbulence with scales smaller than the resolution volume. For VAD scanning, which has the most accurate estimator properties (smallest random error, lowest bias for horizontal wind) according to Koscielny (1984), the estimator efficiency at increasing elevation angle (closer to nadir/zenith) generally decreases for the horizontal wind component, whereas it increases for the vertical wind component. The bias of the horizontal wind component (only a function of the horizontal change of the vertical wind) is constant with elevation angle if symmetric viewing geometries are used (if non-symmetric geometries are used it decreases with increasing elevation angle). The bias of the vertical wind component, due to changing horizontal wind components, decreases with increasing elevation angle. Separating the error of the retrieved wind into variances and bias, Koscielny (1984) states that the estimator variances (random errors in the retrieved parameters) can be reduced by increasing sample size. Biases

however depend on the (unknown) spatial derivatives of the wind field, as well as geometric factors for the sampling techniques. They cannot be reduced by increasing the sample size (for ground-based systems).

An early study by Banakh et al. (1995) analyses the influence of turbulence on wind profile representativeness close to the surface based on an idealized, isotropic representation of the turbulence tensor. Obtaining analytic formulas for the total random wind profiling error, they report a strong dependence of the error on stability (stable, neutral or weakly unstable stratification and the corresponding atmospheric parameters of turbulent intensity, integral correlation scale and wind speed). In agreement with Koscielny (1984), their results show a lowered profile error with decreasing scan elevation (further away from nadir/zenith). Further, using the simplified model they are able to describe the behavior of the error in relation to the number of scanner rotations. As expected, increasing the number of rotations lowers the estimation error, although not below a certain threshold.

A similar approach, using an idealized von-Karman turbulence model, is chosen by Frehlich (2001) in his study of space-based Doppler lidar error. He investigates the representativeness of Doppler lidar radial velocities in-depth. According to Frehlich (2001) the most important source of error is the random error of the wind profile measurement. In-line with Baker et al. (1995), he suggests a division of the total random error into two components: First, the random error caused by variance of the radial velocity measurement due to lidar noise, turbulence in the measurement volume and stability of the wind retrieval algorithm. Second, the representation or sampling error, which is determined by the lidar shot pattern and the random variations of the wind field within that volume. Using idealized conditions (simple forward-aft scan pattern and idealized turbulence), Frehlich (2001) also produces an estimate of the sampling error which appears due to the retrieval of the horizontal wind from multiple radial velocity measurements. Frehlich (2001) concludes noting that for coherent Doppler Lidar the random errors of the radial velocity measurement are usually small and that the major error contribution is due to the sampling errors.

A study by Banakh and Werner (2005) presents the numerical simulation of a coherent Doppler lidar, including the Doppler velocity estimation from spectral analysis. Due to the available computer power at the time their study is limited to simulating random radial velocity fluctuations based on a statistical model. Confirming Banakh et al. (1995), they obtain relative errors on estimates of the dissipation rate, wind velocity variance and outer scale of turbulence as functions of the number of scan rotations. In a second part, they also analyze space-based Doppler lidar retrievals. They find a strong impact of SNR on retrieval quality, suggesting the use of spectral accumulation to increase the reliability of the estimated radial velocities. They also

point the way towards retrieving the turbulent quantities from space-based lidar with acceptable uncertainty margins.

Summarizing, many of the older studies put a focus on the effects of turbulence on system characteristics and performance. The reliability of measured radial velocities under different turbulent intensity conditions and its impact on retrieval quality are well described. The older studies have in common that they determine the representativeness of Doppler lidar measurements based on a statistical description of turbulence and for idealized system set-ups. Little focus is put on the violation of the flow homogeneity assumption for wind profiling at high resolution and the possibility of quality control during the wind retrieval process.

Due to the obvious need for accuracy assessment, measurement-centered studies have also presented theoretical concepts for wind profile reliability assessment and uncertainty estimation based on the measured data (Boccippio, 1995; Leon and Vali, 1998; Holleman, 2005; Damiani et al., 2005; Leon et al., 2006; Cheong et al., 2008; Shenghui et al., 2014; Päschke et al., 2015; Wang et al., 2015; Tian et al., 2015). However, as the input truth is unknown in measurement studies, the proposed methods have mostly remained without verification through theoretical investigation. The studies and their proposed methods are not presented here but discussed in the relevant theory section instead (Sec. 3.6), due to their focus on the mathematical tools available in the retrieval process (e.g. the correlation coefficient, condition number and covariance matrix).

2.4.2. Large eddy simulation based analysis

A valuable link between theory and measurements is provided by the LES-based simulation of Doppler lidar measurements. With increasing computer power, theoretical studies began to use a different approach, not relying on statistical simulation of turbulence and its impact on wind profiling performance. Instead, high resolution model output from LES simulations became available as input for Doppler simulators. The simulated LES wind fields are closer to observed atmospheric conditions (due to the usage of spatially filtered NVS equations, giving a realistic description of variance and covariances, except for the highest spectral frequencies, see Sec. 3.7) and consequently the measurement system performance can be evaluated under conditions closer to reality. The big advantage of LES simulators compared to verification measurements is the fact that the atmospheric input data in the sampling volume is known exactly. Consequently, the representation error (Frehlich, 2001) can be avoided. This knowledge is especially valuable for remote sensing by Doppler instruments, as the applied measurement geometries are almost impossible to represent or even retrace by other measurement systems in real-world

measurements. Consequently, multiple studies using Doppler simulators have been presented in the past.

The first usage of LES data for radar radial velocity estimation is presented by Muschinski et al. (1999). In their study a radar wind profiler, including the signal processing, is simulated based on LES data as atmospheric input. They report that the main drivers of (vertical velocity only) total random error are the meteorological conditions (sampling or representation error), with noise from the measurement process itself (random error) playing a minor role.

Based on the study by Muschinski et al. (1999), a study by Scipión et al. (2008) analyses the wind profile retrieval of five radars focused in one point using an LES-based radar simulator. In follow-up studies by Scipión et al. (2009) and Scipion (2011) wind profile retrievals obtained from Doppler beam swinging (DBS) by one scanning radar are analyzed and evaluated. Compared to the LES input-truth, total random errors in the vertical profiles of the horizontal wind retrieved by DBS are found to be considerable under convective conditions (up to $5\,\mathrm{m\,s^{-1}}$). They attribute the largest error contribution to a horizontal shear of the vertical wind under convective conditions, in agreement with Koscielny (1984). Further, the vertical shear of the horizontal wind is important. Consequently, errors are largest in the lower BL where convection is strongest. Also, errors are larger for lower horizontal wind speeds, when the relative contribution of the vertical wind increases. According to their findings, it is impossible to separate these effects from the real wind estimates based on measurements. They also note that the error contribution cannot be fully explained by the above two contributions alone. Overall, Scipión et al. (2009) and Scipion (2011) state that further research is required to investigate the erroneous retrievals under convective conditions. Their findings are corroborated by an experimental study from Cheong et al. (2008), who use an imaging radar to investigate the effect of flow inhomogeneity on DBS. They show that the homogeneity assumption is fulfilled for 10 minute averaged profiling data, but not for 30 seconds averaging intervals. As ADL aims to provide maximum horizontal resolution, resulting in minimum averaging time, the effects of spatial inhomogeneity are assumed to be non-negligible, especially due to the high elevation angles used.

Stawiarski et al. (2013) use a Doppler lidar simulator to investigate errors from horizontal dual-Doppler measurements. Their study provides the basis for the research presented here, but the results have limited applicability to the problem investigated here. In their study, two lidars are used for simultaneous measurements, consequently the measurement geometry and thereby inversion process are different. The horizontal flow inhomogeneity of the atmosphere can be partially resolved because the two lidars measure in a horizontal plane. As the scans are horizontal, no influence of a changing vertical wind on the retrieved horizontal wind exists.

A DBS sodar simulator, based on LES input data and including an emulation of the measurement process, is presented in Wainwright et al. (2014). The study shows that temporal averaging can reduce the root-mean-squared-error (RMSE) in DBS measurements, albeit not below a threshold level. Highest RMSE values for the vertical profiles of the horizontal wind are found at the BL top with values around $1\,\mathrm{m\,s^{-1}}$, this is attributed to the rapidly changing wind field at this height interval.
In a recent study by Lundquist et al. (2015) the errors in DBS wind profiling due to flow inhomogeneity caused by wind turbine wake are determined using an LES-based lidar simulator, including a wind turbine wake model. They report significant errors of up to 30% for the stream-wise hub-height inflow speed and errors exceeding the actual speed for the retrieved vertical wind component.
Another study by Klaas et al. (2015) investigates lidar error in moderately complex terrain. They compare DBS lidar measurements to that of a co-located mast and recreate the observed errors using atmospheric flow models similar to LES. Based on an existing system they take the lidar setup as given and mostly investigate directional dependence of lidar error in complex terrain. They analyze only 10-min intervals and exclude wind speeds $< 4\,\mathrm{m\,s^{-1}}$ from their analysis. Overall, Klaas et al. (2015) report lidar errors between -4% underestimation and $+2.5$% overestimation depending on wind direction.

2.4.3. Existing airborne Doppler simulators

The possibility of an LES-based aircraft measurement simulation was already devised 20 years ago by Schröter et al. (2000) and has since been applied more often (Sühring and Raasch, 2013; Sühring et al., 2019). However, the referred studies focus on the simulation of in-situ sensors to validate aircraft measured turbulent fluxes. Similar to the ideas presented in this work, the studies also use the LES-based simulation to evaluate the reliability of the uncertainty estimation methods (Lenschow and Stankov, 1986; Lenschow et al., 1994) associated with the measurements. Another type of simulation approach is presented by Vihma and Kottmeier (2000). They use coarser resolution output from a meso-scale NWP model for the optimization of airborne sampling strategies. Thereby, they are able to calculate optimal flight pattern for different tasks such as air mass characterization and turbulent flux measurements. Airborne Doppler simulators based on modeled, idealized flow fields have been introduced for radar and lidar systems throughout the last decades. While none of the airborne Doppler simulators utilizes high resolution (O(100 m)) atmospheric input fields, their evolution and findings mandate discussion nevertheless.

Theoretical investigations using measurements conducted in simulated wind fields began simultaneously with the availability of ADL systems for experimental studies (Reitebuch et al., 2001; Leike et al., 2001). However, the first studies had to rely on coarse model output with fully parametrized turbulence (e.g. 7 km grid spacing in Reitebuch et al., 2001, approx. 100 km grid spacing in Leike et al., 2001). Therefore, Leike et al. (2001) adopted the approach by Banakh et al. (1995) with the simulation of idealized isotropic turbulence in the two-dimensional measurement plane based on a von-Karman model. Consequently, the first simulator studies were unable to investigate the wind profiling error induced by BL turbulence in-depth. However, they conducted a full simulation of the radial velocity measurement process, leading to valuable insight into its reliability and noise level. Hence, the behavior of the radial velocity measurement is well documented, it is shown to introduce negligible error for this work in Sec. 6.2.III.

The first ADR simulator with an order of magnitude higher resolution atmospheric input data, based on convection permitting NWP simulations, is presented by Lorsolo et al. (2013) for an ADR flying in a hurricane wind field on the NOAA WP-P3D. They analyze the retrieval quality of the Doppler processing software (which was treated as a grey box previously) for an idealized ADR free of system error. Hence, they are able to attribute all observed retrieval error to the retrieval algorithm. They use a single time step of model output from an NWP model with a grid spacing of 1.5 km as wind field input for the simulation of radial velocities. Generally good performance of the retrieved wind speed with a RMSE $< 2\,\mathrm{m\,s^{-1}}$ compared to the input truth is reported. However, they also find small but statistically significant bias of the retrieval. A limitation of the Lorsolo et al. (2013) study is the still coarse NWP model resolution. As their model is only convection permitting it does not allow for direct simulation of BL turbulence. Hence, it is not expected that their study captures the full retrieval error due to flow inhomogeneity.

In another simulator study for ADR wind profiling, Guimond et al. (2014) analyze the retrieval accuracy for the IWRAP and HIWRAP ADRs which are based on high and fast flying long-range aircraft. Due to very rapid scanning characteristics (60 rotations per minute) they acquire multiple radial velocity samples along and across the flight track for a given volume. Thereby, it is possible to extract the horizontal wind distribution using a volume-based inversion process. They analyze the effect of the airborne radar geometry and radar noise on the retrieved wind speed and report errors of $7-8\%$ for hurricanes. Further, they are able to show that due to their viewing geometry the errors have a strong across-track dependence. Similar to the Lorsolo et al. (2013) study, a limitation of Guimond et al. (2014) is the coarse grid spacing of the wind field

input data (2 km horizontally and 0.65 km vertically), which does not allow for adequate BL turbulence development.

The same group of authors published another study (Didlake et al., 2015) with a similar system and model setup but for a coplane dual-Doppler technique retrieval strategy. A simplified version of the coplane technique is also analyzed through ADR simulations by Melhauser and Zhang (2016). Last, another interesting approach is developed by Bousquet et al. (2016). They simulate ADR radial velocity measurements not based on model output, but based on a wind field retrieved by synthesis of volumetric scans from an array of ground-based radars. Thereby, they are able to determine the agreement between observation systems in both directions.

Summarizing the previous sections, newer Doppler simulator studies often utilizes LES generated wind fields for a more realistic description of radial velocity estimation and wind retrieval properties under inhomogeneous flow conditions. For ground-based systems a number of investigations detailing the problems associated with wind profiling exist. For airborne systems, studies by Reitebuch et al. (2001); Lorsolo et al. (2013); Guimond et al. (2014) and Didlake et al. (2015) have shown the importance of simulator studies using coarser resolution model output and focused on errors introduced due to the measurement system deficiencies. A simulation of an airborne Doppler wind profiling system with complex scanning geometries based on high resolution atmospheric wind fields (O(100 m)) was missing prior to the work presented here. Thereby, as discussed above, all simulator studies so far are unable to capture the effect of flow inhomogeneity due to BL turbulence on the retrieval accuracy, a problem which is overcome by this work.

3. Concept and theory of airborne Doppler lidar measurements

In this chapter the concept and theory of the used measurement system and retrieval techniques are detailed. The discussion includes a description of the potential errors impacting measurement system and retrieval accuracy, both for wind profiling and nadir measurements. The new ADL system is intended for wind profiling at highest resolution in the BL, consequently the effect of flow inhomogeneity on retrieval error needs to be considered and is discussed. To this end, the existing theory on covariance matrix based uncertainty propagation is extended and a data-driven uncertainty estimation method for usage in AVAD wind profiling is proposed, which is then validated using the ADLS in the next chapters. Finally, insight into the theoretical foundation of the used LES model is provided, as it forms the basis of the ADLS.

3.1. Doppler lidar measurement principle

Active remote sensing by means of lidar (an acronym for LIght Detection And Ranging) utilizes electromagnetic waves in a range from 250 nm to 11 μm. Since the availability of laser technology as an emitting source in the 1960s, lidar technology has evolved rapidly from its first mentioning in literature (Middleton and Spilhaus, 1953). Today, a variety of different lidar technologies exists, all are based on the principle of detecting the return signal of emitted radiation. Depending on the process responsible for the return signal, different types of lidar exist. As outlined in Weitkamp (2005), among them are (the characteristic process is given in brackets): elastic backscatter lidar (Rayleigh and Mie scattering), differential absorption lidar (difference in absorption coefficients for different wavelengths), Raman lidar (inelastic Raman scattering), fluorescence lidar (stimulation for resonance) and the Doppler lidar (frequency shift of the return signal due to the movement of the scatterers).

The laser as a light source

The basis of modern lidar systems is the emission of laser beams into the atmosphere, the return signal of which is then detected and analyzed. A laser (Light Amplification by Stimulated Emission of Radiation) is a light source which can create a focused beam of coherent and

monochromatic light. The working principle of a laser is based in quantum mechanics, in which a cascading effect of stimulated photon emission is used to create the light pulse.
To do so, lasers contain an active medium in which a population inversion of photons can be created through an energy pump, e.g. high energy light flashes. Thereby, the number of photons at a higher, excited (discrete) energy level is increased compared to thermal equilibrium. Transition of a photon to lower energy levels leads to emission of light with a defined wavelength, corresponding to the energy difference between the two discrete energy levels (Demtröder, 2006):

$$\Delta E = E_2 - E_1 = \hbar\omega. \tag{3.1}$$

Transition can happen spontaneously or by stimulation through another photon with matching wavelength (energy) for the transition. In a laser, the stimulated emission is intensified by keeping the photons in an active medium using the mirrors of an optical resonator (while pumping energy and thereby supporting the population inversion). The optical resonator also serves to limit the trapped energy to desired modes. If the number of photons available for stimulated emission (at excited energy levels) surpasses the losses due to absorption, scattering and diffraction of photons in the active medium, stimulated emission occurs in a cascading effect at one specific wavelength. One of the resonator mirrors is semi-transparent and thereby allows for emission of the coherent, monochromatic light.
The WTX system used in this work is a pulsed solid-state laser in which the active medium is an erbium doped yttrium aluminum garnet (Er:YAG). The laser emits coherent radiation at a wavelength of 1617 nm with a pulse repetition frequency of 750 Hz and a typical pulse energy of 2.7 mJ. The temporal envelope of the Gaussian pulse is 1.8×10^{-7} s (1/e intensity duration, full width at half maximum (FWHM) duration is 3×10^{-7} s). The laser beam is widened by a telescope aperture to a transverse beam diameter of 10 cm (1/e intensity radius). Thus, eye-safety is ensured for unaided viewing and a staring beam, which is upgraded to include aided viewing for a moving beam (laser class 1M).

The atmospheric return signal

The lidar used in this work is a pulsed coherent (heterodyne-detection) Doppler lidar. For pulsed lidar, range resolved information can be obtained because the time difference between the laser pulse emission and the detection of the returned signal is a measure of the distance of

the scatterers along the beam. The power of the returned signal P from a distance r along the beam can be expressed through the lidar equation (e.g. Wandlinger, 2005):

$$P(r,\lambda) = P_0 \eta A \frac{c\Delta T}{2} \beta(r,\lambda) \frac{1}{r^2} e^{-2\int_0^r \alpha(r',\lambda)dr'}. \tag{3.2}$$

It is thereby dependent on the emission power of the laser source P_0, a lidar detection efficiency parameter η, the detector area A and the spatial effective pulse length $\frac{c\Delta T}{2}$. Additionally, the ability of the atmosphere to generate backscatter signal also varies, it is contained in the term of the backscatter coefficient β, which is made up of the scatterer concentration and their scattering cross-section. Generally, the received signal power decreases with $\frac{1}{r^2}$ for increasing range due to geometric reasons (a factor 4π is contained in both the backscatter and range term and thereby cancels out). Additional losses can occur during signal transmission through the atmosphere due to extinction α, which consists of absorption and scattering of radiation (the factor 2 is due to the two-way transmission path). All scattering and extinction processes are functions of the signal wavelength λ.

Although the backscatter coefficient contains contributions by both molecules and aerosols, the Doppler lidar used in this work relies strongly on the backscatter produced by aerosols, due to its wavelength of 1612 nm (Mie scattering regime). Hence, the Doppler lidar sensed radial velocity is the average aerosol movement in the illuminated volume. Scattering by aerosols has the advantage of a more defined Doppler shift (narrow spectral peak), due to their higher mass and thereby reduced random motion compared to molecules. Downward observing airborne lidar measurements have an advantage over ground-based systems, because the concentration of scatterers increases with distance, as aerosol loads typically increase towards the ground. Thereby, the signal losses due to increasing distance and increasing transmission losses are partially compensated by an increasing $\beta(r,\lambda)$, leading to improved return signal strength.

Wind measurements

Doppler lidar exploits the frequency shift in the returned electromagnetic radiation when a scattering process occurs on a moving scatterer (the so-called Doppler shift). The Doppler shift Δf is proportional to the projection of the scatterer velocity in beam direction v_r (Werner, 2005):

$$\Delta f = \frac{-2v_r}{\lambda_L}. \tag{3.3}$$

The above formulation implies that motion towards the lidar is negative. The observed Doppler shifts are extremely small for atmospheric wind speeds. For example, a wind speed of v_r =

$1\,\mathrm{m\,s^{-1}}$ results in a frequency shift of only $\Delta f = 1.24\,\mathrm{MHz}$ for a laser wavelength of 1617 nm, as in the system used in this work. The WTX detector samples with a bandwidth of 95 MHz, which corresponds to a maximum retrievable radial velocity of $\pm 38\,\mathrm{m\,s^{-1}}$.

Signal processing

Due to the small absolute frequency difference of the Doppler shift a very stable laser frequency is required to enable measurement of the small changes. In order to capture these small frequency changes heterodyne detection is employed, in which the returned signal is mixed with the reference signal from a local-oscillator of known frequency (Werner, 2005). As the transmitted signal frequency is also monitored using the same local-oscillator the Doppler shift can be accurately determined despite the small difference.

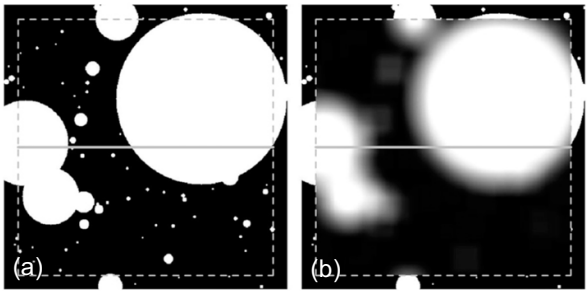

Fig. 3.1.: Illustration of the PVA-effect from Brugger et al. (2016). (a) A true field of white circles on a black background. (b) The same field after spatial averaging with a 4% area averaging kernel. The black-white variance along the gray line is reduced by 20% and the integral length scale is increased by 18%. © American Meteorological Society. Used with permission.

The returned signal is recorded at a sampling rate (f_{spl}) of 250 MHz, which corresponds to an along-beam spacing of 0.6 m for the round trip distance. The recorded signal is then subjected to a fast Fourier transform (FFT) in order to identify the frequency difference of the Doppler shift and thereby the radial velocity in so called range gates. A variable number of samples per range gate (SpG) can be input into the FFT used for estimation of the radial velocity. In this work a value of 116 is used, corresponding to a range gate length of

$$\Delta p = \frac{SpG \cdot c}{2 f_{spl}} = 69.53\,\mathrm{m}. \tag{3.4}$$

The range gates are chosen to be non-overlapping. In order to reduce the random error in the estimation process each radial velocity is computed by averaging the spectra of 75 pulses, which gives a data out rate of 10 Hz at the WTX PRF of 750 Hz. After estimation of the spectral peak, the WTX system calculates a wideband SNR on a 6 MHz bandwidth interval by comparing the signal peak to the noise floor for every range gate. The SNR is used for quality control of the acquired data in Sec. 6.2.III.

Due to the characteristics of the Gaussian pulse, and the radial velocity estimation process through FFT processing in range gates, the estimated radial velocity is an average of the parti-

cles illuminated by the pulse as it moves through the range gate volume. This so-called pulse volume averaging constitutes a difference of lidar measurements compared to in-situ point measurements and needs to be accounted for, especially when calculating variances (Fig. 3.1). Frehlich (1997) and Frehlich et al. (1998) provide a mathematical model for the estimation of the radial velocity v_D at range gate center distance R_{dis} as

$$v_D(R_{dis},t) = \frac{1}{\Delta p}\int_{R_{dis}-\Delta p/2}^{R_{dis}+\Delta p/2} v_{pulse}(r,t)dr. \tag{3.5}$$

In this, v_{pulse} is the average weighted velocity given by a pulse centered at range r, it is given as

$$v_{pulse}(r,t) = \int_{-\infty}^{\infty} v_r(x,t)I_n(r-x)dx. \tag{3.6}$$

Thereby, the estimated velocity is a convolution of the pulse-weighted radial velocity with the range gate weighting function. The weighting function of the pulse is given by its Gaussian envelope,

$$I_n(r) = \frac{2}{\sqrt{\pi}\sigma_p c}e^{-4r^2/(\sigma_p^2 c^2)}, \tag{3.7}$$

with σ_p being the $1/e$ width of the pulse, which is 1.8×10^{-7} s in case of the WTX. This corresponds to a spatial full-width at half maximum of

$$\Delta r = \sqrt{ln2}c\sigma_p = 45\,\mathrm{m}. \tag{3.8}$$

3.2. Airborne Doppler wind measurements

For airborne measurements, the lidar does not sense solely the atmospheric wind contribution. Instead, the combination of the aircraft and atmospheric motion vectors gives the measured radial Doppler velocities. Typically, the aircraft motion vector contribution is an order of magnitude larger than the atmospheric contribution. It thereby needs to be determined with high accuracy in order to allow for atmospheric wind measurements.

Conveniently, the two velocities are treated separately by defining two separate coordinate systems (Fig. 3.2), following the theory outlined in Lenschow (1972) and Leon and Vali (1998). The wind field is typically based in an earthbound coordinate system (superscript E) oriented east-north-up (ENU). The aircraft and associated components are typically based in an aircraft related coordinate system (superscript A), which is oriented along aircraft-right wing-down (ARD). To transfer between the two systems, coordinate transformations are required, the details are given in Appendix A.2. Additional coordinate transforms are necessary in real-world

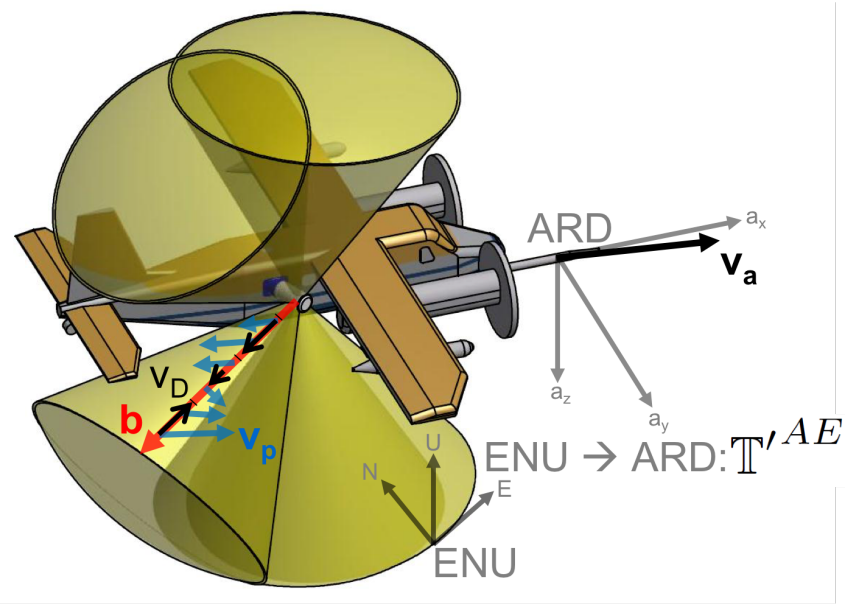

Fig. 3.2.: Schematic depiction of the ADL system simulated in the ADLS. The sketch is based on an upcoming system for the Dornier 128-6 aircraft of the TU Brunswick (D-IBUF, Corsmeier et al., 2001), originally intended for usage in the measurements conducted in this work. The lidar is inside the aircraft 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 equation 3.11.

ADL measurements in order to locate the measurements in a geodetic coordinate system (GEO), using an intermediate Earth-centered, Earth-fixed (ECEF) coordinate system. The GEO and ECEF transformations are also detailed in the Appendix A.2.

The radial Doppler velocity v_D 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 are projected onto the beam direction vector $\boldsymbol{b}$, which is defined as a unit vector:

$$v_D = \boldsymbol{b}^A \cdot \left(\boldsymbol{v}_p^A + \boldsymbol{v}_L^A \right). \tag{3.9}$$

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

$$v_D = \boldsymbol{b}^A \cdot \left(\boldsymbol{v}_p^A - \boldsymbol{v}_a^A - \boldsymbol{\omega}^A \times \boldsymbol{r}^A \right). \tag{3.10}$$

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 ADL system measures foremost the

aircraft velocity, as it presents a vector with very large magnitude compared to the atmospheric motion vector and the aircraft moment arm. For the ADLS, 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 Appendix A.2). 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:

$$v_D = \boldsymbol{b}^A \cdot \left(\mathbb{T}'^{AE} \boldsymbol{v}_p^E - \mathbb{T}'^{AE} \boldsymbol{v}_a^E - \boldsymbol{\omega}^A \times \boldsymbol{r}^A \right). \tag{3.11}$$

In order to distill the atmospheric contribution from the measurement, the contributions of the aircraft motion and aircraft moment arm to the measured radial velocity are removed in a process termed motion correction:

$$v_{COR} = v_D + \boldsymbol{b}^A \cdot \left(\mathbb{T}'^{AE} \boldsymbol{v}_a^E + \boldsymbol{\omega}^A \times \boldsymbol{r}^A \right). \tag{3.12}$$

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).

3.3. Wind profiling for mean flow observation

An important task of existing ADL systems is wind profiling to obtain the mean flow state of the atmosphere in the retrieval volume, e.g. using an AVAD retrieval procedure (Sec. 2.3). In the AVAD technique, multiple radial velocity measurements v_{COR_n} under different beam pointing directions $\boldsymbol{b}_n$ are sampled from an atmospheric wind field with mean wind vector $\boldsymbol{v}_p$:

$$\begin{bmatrix} v_{COR_1} \\ v_{COR_2} \\ v_{COR_3} \\ \vdots \\ v_{COR_n} \end{bmatrix} = \begin{bmatrix} b_{x_1} & b_{y_1} & b_{z_1} \\ b_{x_2} & b_{y_1} & b_{z_1} \\ b_{x_3} & b_{y_1} & b_{z_1} \\ \vdots & \vdots & \vdots \\ b_{x_n} & b_{y_n} & b_{z_n} \end{bmatrix} \begin{bmatrix} v_{p_x} \\ v_{p_y} \\ v_{p_z} \end{bmatrix}. \tag{3.13}$$

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

$$v_{COR} = \mathbb{G} \boldsymbol{v}_p. \tag{3.14}$$

Knowing v_{COR} and $\mathbb{G}$, the inverse problem is then solved by calculating the inverse $\mathbb{G}^{-g}$ of the beam pointing matrix $\mathbb{G}$, in order to retrieve an estimate of the wind vector $\boldsymbol{v}^R$ responsible for the observations:

$$\boldsymbol{v}^R = \mathbb{G}^{-g} v_{COR}. \tag{3.15}$$

From the retrieved wind vector $\boldsymbol{v}^R = [u, v, w]$ the wind speed is calculated as $V_m = \sqrt{u^2 + v^2 + w^2}$. The general inverse $\mathbb{G}^{-g}$ of the over-determined least-squares problem can be calculated as (Menke, 2012):

$$\mathbb{G}^{-g} = [\mathbb{G}'\mathbb{G}]^{-1}\mathbb{G}'. \tag{3.16}$$

Equation (3.16) presents a least-squares (LSQ) solution to the problem. It should be noted that although the airborne retrieval procedure is called AVAD, it is more similar to a VVP technique (Waldteufel and Corbin, 1978) than to the original VAD technique (Browning and Wexler, 1968). The underlying basic functions are not inherently orthogonal for AVAD (and VVP), whereas they are for VAD due to the Fourier decomposition of the wind field. Instead of calculating the general inverse in the above way, in this work a singular value decomposition (SVD) is performed, which yields more insight into the numerical stability of the inversion process compared to the above direct solution (Boccippio, 1995):

$$\mathbb{G} = \mathbb{U}\mathbb{S}\mathbb{W}', \tag{3.17}$$

$$\mathbb{G}^{-g} = \mathbb{W}_p\mathbb{S}_p^{-1}\mathbb{U}_p'. \tag{3.18}$$

SVD results in an orthogonal decomposition, $\mathbb{U}'\mathbb{U} = \mathbb{W}'\mathbb{W} = \mathbb{I}$, 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:

$$CN = \frac{max(\lambda_S)}{min(\lambda_S)}. \tag{3.19}$$

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

$$v_{COR}^R = \mathbb{G}\boldsymbol{v}^R. \tag{3.20}$$

This average, reconstructed radial velocity differs from the measured radial velocities due to flow inhomogeneity smaller than the measurement volume as well as model mis-specification (as well as other factors such as measurement system noise not considered here). The differences between the average and measured radial velocities is denoted as the residuals in the following:

$$v_{COR}^{res} = v_{COR} - v_{COR}^{R}. \tag{3.21}$$

Using the average reconstructed radial velocity, the R^2 is defined as

$$R^2 = 1 - \frac{\sum_n (v_{COR_n} - v_{COR_n}^{R})^2}{\sum_n (v_{COR_n} - \sum_n v_{COR_n})^2}. \tag{3.22}$$

Assessing the reliability of the parameters retrieved through the inversion process is a common problem in inverse theory and two metrics are investigated as a part of this work. The coefficient of determination R^2 is often used to detect a violation of the flow homogeneity assumption for ground-based measurements (Päschke et al., 2015; Wang et al., 2015). When using this approach, it is assumed that deviations from homogeneous flow conditions show up as deviations of the measurements from the LSQ-fit. Therefore, it is commonly assumed that the R^2 captures the degree of violation of the homogeneity assumption. Often used quality control 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 of the vertical wind in the horizontal direction 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, as only flow inhomogeneity smaller than the scan volume size decreases the R^2. In Sec. 5.1.2 further complications with using R^2 as a quality control criteria are revealed, and results from the work presented here suggest that quality control with R^2 is strongly not advisable for ADL measurements.

If the matrix inversion is performed based on a singular value decomposition, additional quality control criteria such as the condition number CN (Eq. 3.19), 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 as a result the inferred

system state is prone to a large error amplification (Boccippio, 1995). The CN is often used as a quality control 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). The functionality of CN-based quality control is shown to be robust and reliable in Sec. 5.1.3.

3.4. Nadir for high resolution vertical wind observation

Another important task of ADL systems is the high-resolution investigation of the vertical wind by means of nadir measurements (Kiemle et al., 2011; Chouza et al., 2016b; Witschas et al., 2017). Extending the research conducted by Kiemle et al. (2011), in this work, ADL measurements are used to retrieve the first and second order moments and their associated uncertainties for the vertical wind in a turbulent BL for the first time.
From here on, this work combines the notation of Lenschow et al. (1994, 2000) and Strauss et al. (2015) for the description of statistical moments and their associated uncertainties. One would like to know the ensemble mean $\overline{\mu^1} = \langle q(t) \rangle$ of an ergodic (and thereby stationary) time series $q(t)$. As this ensemble mean is impossible to determine with current measurement systems, the ensemble mean is estimated as the time average,

$$\overline{\mu^1}(T) = \frac{1}{T} \int_{-T/2}^{T/2} q(t) dt. \tag{3.23}$$

This replacement introduces random differences for the individual realizations, but for the ensemble mean $\langle \overline{\mu^1}(T) \rangle = \langle q(t) \rangle$ holds. With the assumption of ergodicity, ensemble mean and time average converge for long measurement durations, $\lim_{T \to \infty} (\overline{\mu^1}(T)) = \langle q(t) \rangle$.
The autocovariance function of the time series at lag τ is defined as

$$\mathcal{R}(\tau) = \langle [q(t) - \langle q \rangle][q(t+\tau) - \langle q \rangle] \rangle, \tag{3.24}$$

which can be converted to the autocorrelation function ρ through

$$\rho(\tau) = \frac{\mathcal{R}(\tau)}{\overline{\mu^2}}. \tag{3.25}$$

From the autocorrelation function the integral scale $\mathcal{T}$ can be calculated as

$$\mathcal{T} = \int_0^\infty \rho(t) dt. \tag{3.26}$$

The integral scale is a characteristic measure of the time it takes for the time series to become fully de-correlated from itself (e.g. when 'memory' is lost).
Similarly to the mean, the second-order moment $\overline{\mu^2}$ is defined as

$$\overline{\mu^2} = \langle (q(t) - \langle q(t) \rangle)^2 \rangle = \langle q(t)'^2 \rangle, \tag{3.27}$$

where the prime denotes deviations from the ensemble average $\langle q(t) \rangle$. For its implementation in real-world experiments

$$\overline{\mu^2}(T) = \frac{1}{T} \int_{-T/2}^{T/2} [q(t) - \overline{\mu^1}(T)]^2 dt \tag{3.28}$$

is used. The replacement introduces systematic and again random errors, the associated uncertainty estimation is discussed in the relevant section dealing with statistical errors due to the finite time series length (Sec. 6.2.VIII and Appendix A.9).
Besides the calculation of the mean quantities, an analysis of the time series using spectral techniques is also interesting (Sec. 2.1). The spectra provides a decomposition of the variance contributions by frequency f (or equivalently wavenumber) so that

$$\overline{\mu^2} = \int_0^\infty S(f) df. \tag{3.29}$$

In practice, the discrete spectrum of a time series is calculated through an FFT of a finite time series. In the following, the term spectrum refers to the squared absolute value of the discrete complex one-sided FFT, which is used for practical implementation in measurements. The lidar data is finite, further it can exhibit data gaps and other undesired effects which are not captured in the theoretical description of the spectral analysis detailed here. Therefore, in order to distill spectra from the measured time series a number of data conditioning techniques have to be applied in order to obtain reliable and meaningful results, these are discussed in Sec. 6.2.VIII.
By definition, the integration over the spectrum must yield the overall variance of the time series. Similarly, the level of the spectral coefficient for the frequency bins is related to the variance contribution at the respective frequency (although not directly proportional, depending on the chosen display style). Further, based on theoretical assumptions, spectra are expected to exhibit typical shapes. Thereby, through agreement or disagreement of the observed spectra, insight into the structure of the atmospheric turbulence can be gained.
The spectra can be linked to the autocovariance function of Eq. 3.24 through the
Wiener–Khinchin theorem (see Taylor, 1938), which means that both forms contain the same

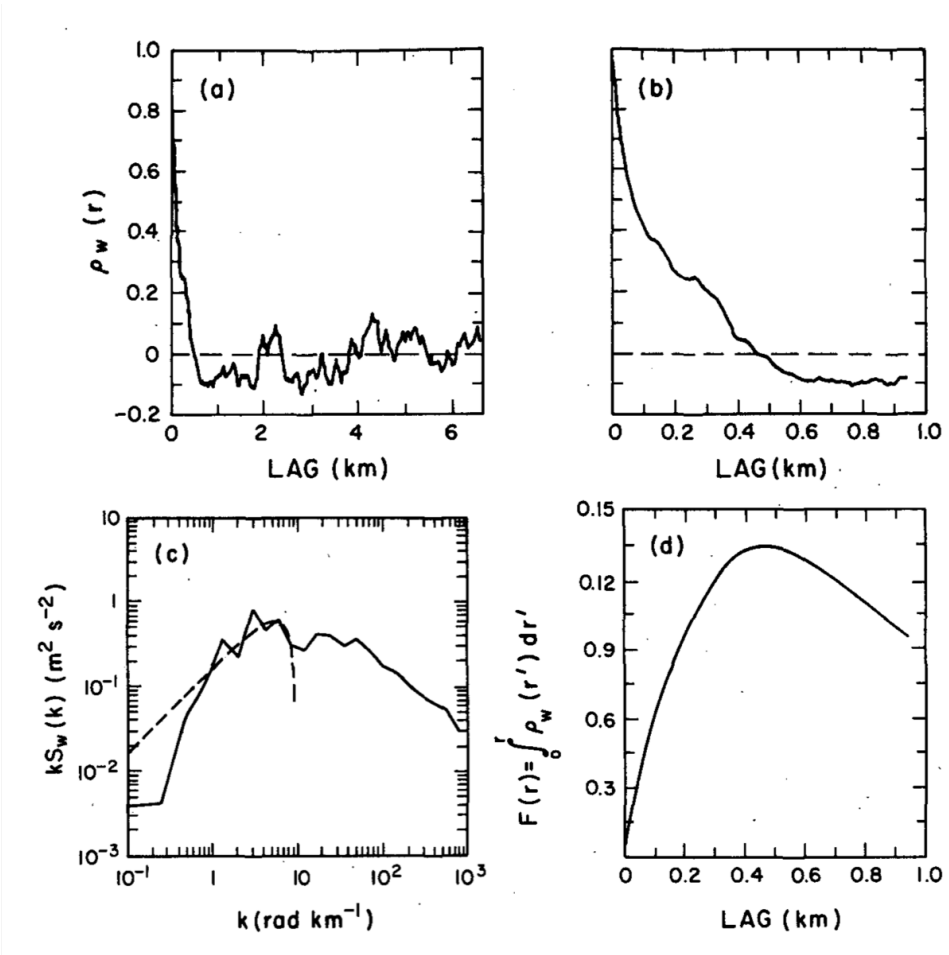

Fig. 3.3.: An example on autocorrelation, associated spectrum and integral scale from aircraft measured vertical wind time series, as presented by Lenschow and Stankov (1986). Shown are: (a) Typical autocorrelation function with random variations beyond the first zero crossing. (b) Zoom into the correlated region with rapid decay of the correlation. (c) The associated spectrum. (d) The integral used for estimation of the integral length scale. © American Meteorological Society. Used with permission.

amount of information. Consequently, the integral scale (Eq. 3.26) can also be calculated through a fit to the spectra. However, in this work the more common calculation through the autocorrelation function by integration of the autocorrelation function up to the first zero-crossing is used for experimental implementation (Lenschow and Stankov, 1986):

$$\mathcal{T} = \int_0^{\rho=0} \rho(t)dt. \tag{3.30}$$

Using the calculation through the autocorrelation has the benefit of a more direct and stable uncorrelated noise removal approach (Lenschow et al., 2000; Frehlich and Cornman, 2002). The temporal integral scale $\mathcal{T}$ can also be converted to a length scale $\mathcal{L}$ through application of

Taylor's hypothesis. The integral length scale is an another important parameter often used for describing the properties of turbulent flows. It denotes the distance up to which the turbulent structures are correlated with themselves, it is thereby similar to a 'memory' of the flow.

3.5. Measurement platform errors

In the ADLS a so-called ideal system is assumed, without error in any of the system components consisting of aircraft, scanner and lidar. Thereby, the motion-corrected Doppler velocity is equal to the projection of the atmospheric wind vector on the beam vector itself, as all other velocity contributions are known exactly (Eq. 3.12). The real-world ADL measurements differ from the idealized ADLS measurements in important ways. Because the state of the real-world ADL measurement system is imperfectly known, errors can be associated with any of the terms in the motion correction process.

Using the motion corrected radial velocities for wind profiling can introduce further error. As no system errors are present in the ADLS, only the violation of the flow homogeneity assumption used for AVAD retrieval impacts the wind profiling error in the ADLS. Thereby, the error in wind profiling due to the violation of the homogeneity assumption can be extracted for the first time and is discussed in detail in the next section. The retrieval error is equally present in real-world ADL measurements but in general inaccessible, for this reason a data-driven uncertainty estimation method is developed using the ADLS (for the definition of the terms error, uncertainty and accuracy as used here see Appendix A.1).

Additional sources of error exist in the estimation of statistical quantities from nadir measurements (e.g. the first and second order moment of the vertical wind) by means of ADL due to the measurement platform characteristics. Nadir retrieval of statistical quantities is only conducted for illustrative purpose in the ADLS because the used LES simulations do not represent the turbulent wind field in enough detail for systematic analysis. Due to the used LES grid spacing of $\Delta x = 10\,\mathrm{m}$, the LES does not fully represent the turbulent spectrum below O(100 m) (Sec. 3.7). As this region is important for retrieval of vertical wind statistics and also problematic for the lidar, due to effects such as PVA, the LES cannot be used for reference. Hence, systematic analysis of nadir measurement retrieval accuracy is solely conducted for the real-world ADL system.

The real-world ADL measurement platform is much more complex than the idealized platform in the ADLS, therefore a number of synchronization, calibration and correction procedures have to be implemented in order to achieve high measurement quality. Corresponding with these procedures, a number of additional sources of error and associated uncertainties (both

systematic and random) are evaluated to control the performance of the applied techniques. In order to achieve real-world data synchronization, calibration and correction I follow the approach of existing studies, which are combined, refined and extended as a part of this work. Overall, the terminology and methodology of uncertainty estimation presented by Damiani and Haimov (2006) and Strauss et al. (2015) is followed here (who developed their characterizations for ADR), but with important changes and extensions described below. As a result, to my knowledge, the most comprehensive accuracy assessment of real-world ADL data is presented in the following. The uncertainties are distinguished by their generating source of error for the first-order moment (systematic offset $\overline{\beta^1}$, random uncertainty $\overline{\varepsilon^1}$ in $\mathrm{m\,s^{-1}}$) and the second-order moment (systematic offset $\overline{\beta^2}$, random uncertainty $\overline{\varepsilon^2}$ in $\mathrm{m^2\,s^{-2}}$). Effects not simulated in the idealized ADLS, but important for both wind profiling and nadir vertical wind retrievals in real-world ADL measurements, are:

(I) A time synchronization between aircraft and lidar has to be achieved, taking into account the different temporal resolutions of the measurement systems. In this work, time synchronization is achieved based on a new procedure developed as a part of this work, which exploits the changing ground return distance sensed by the lidar during aircraft maneuvers. The procedure enables high accuracy time synchronization, the remaining uncertainties are denoted as β_t and ε_t for systematic and random effects, respectively (separately for first- and second-order moments).

(II) The static beam pointing-angle of the lidar mounting has to be determined in a procedure called beam pointing-angle calibration. In this work, beam pointing-angle calibration is achieved using techniques outlined in Haimov and Rodi (2013). Uncertainties associated with the knowledge of the static beam pointing-angle are denoted as β_b for systematic contributions and ε_b for random contributions (separately for first- and second-order moments).

(III) The lidar signal quality is dependent on the return signal strength, which has to be evaluated and quality controlled. In the high signal return power regime, an uncorrelated random error is present in the lidar measured velocities due to random fluctuations of the lidar signal, which impacts the measured atmospheric variance (Frehlich et al., 1994). In this work, the SNR-dependent systematic offset $\overline{\beta^2_{rv}}$ due to radial velocity noise is estimated using the first-lag autocovariance method proposed by Lenschow and Kristensen (1984), Mayor et al. (1997) and Lenschow et al. (2000), which was shown to be a reliable estimator by Frehlich and Cornman (2002). Random uncertainties introduced by the

estimation of the systematic offset are denoted as $\overline{\varepsilon_{rv}^2}$. The lidar signal quality control procedure also encompasses SNR filtering as well as cloud and ground return identification.

(IV) Finally, a motion correction of the measured lidar radial velocities has to be conducted. A two step procedure is presented here: First, the navigational data available from the aircraft INS is used for initial motion correction of all lidar data. The accuracy of this correction depends on the aircraft motion and attitude measurement accuracy. Going beyond Strauss et al. (2015), based on Godwin et al. (2012) and in-line with Tucker et al. (2018), Lux et al. (2018) as well as Ellis et al. (2019) (for ADR), it is demonstrated that the exploitation of the measured ground return velocities allows for a quantification of errors associated with the INS-based aircraft motion correction (as well as a distinction between flight state). In a second step, it is shown that the measured ground return velocities can then also be used for a refined motion correction, which reduces the motion correction error and provides strongly improved measurement accuracy compared to solely relying on the INS. The ground return based procedure enables aircraft motion and attitude accuracy assessment and correction also for strongly non-steady aircraft motion states. Remaining systematic uncertainties associated with the knowledge of aircraft motion are denoted as β_{ac} and random uncertainties as ε_{ac} (separately for first and second-order moments).

Quantification of the uncertainties (I-IV) is sufficient to assess the accuracy of the ADL measured radial velocities. However, if a retrieval of the atmospheric wind components is required, a number of additional effects require consideration:

(V) Instead of directly sensing the wind velocity, the lidar measures the velocity of aerosol transported by the wind. The aerosol velocity can differ from the wind velocity (e.g. due to a settling of the aerosol towards the ground). Generally, the terminal settling velocity and random motion of the scattering particles at $1.6\,\mu$m wavelength are very small (Werner, 2005; Seinfeld and Pandis, 2006). Contrary to ADR, the terminal settling velocity is thereby not considered to be an important source of systematic offset β_{v_t} or random uncertainty ε_{v_t} for ADL measurements of first- or second-order moments.

Quantification of the uncertainties (I-V) is sufficient to assess the accuracy of the ADL measured atmospheric velocities utilized for wind profiling of the horizontal wind. Additionally, however, the wind profiling accuracy is also dependent on assumptions made in the AVAD retrieval, which can be investigated using the ADLS and are discussed in the next section.

ADL measurements of the vertical wind, and its associated statistical parameters such as the mean, variance and integral length scale, are subject to an additional number of effects requiring consideration:

(VI) Due to the non-steady aircraft motion and attitude the lidar beam does not point nadir at all times. The ADL thereby does not measure the vertical wind component only but is contaminated by an additional horizontal wind contribution instead. The unwanted contribution is removed based on profiles of the horizontal wind retrieved in the vicinity of the nadir legs, as is done by Chouza et al. (2016b). This procedure introduces the systematic uncertainty $\overline{\beta^1_{HC}}$ and the random uncertainty $\overline{\varepsilon^1_{HC}}$ for the first-order moment. The error in the horizontal wind removal can be bounded conservatively by a comparison of the leg-averaged vertical wind mean to the reference value of $0\,\mathrm{m\,s^{-1}}$ mandated by theory for continuity reasons. For the vertical wind variance, the non-nadir pointing introduces an additional source of variance due to the projection of the horizontal wind into the measurement. Both, the systematic effect of cross-contamination due to variations of pitch and roll $\overline{\beta^2_{HC}}$, as well as the random effect of cross-contamination due to non-zero mean pitch and roll $\overline{\varepsilon^2_{HC}}$ are based on Strauss et al. (2015). Their theory and analysis is extended to account for the reduced cross-contamination effect made possible by the horizontal wind removal.

(VII) Due to the spatial extent of the lidar pulse, the measurements are not conducted at individual points in time and space. Rather, the measurements represent averages over a pencil-shaped volume instead, thereby causing the PVA effect. Consequently, the sampled variance is altered (usually reduced) compared to an in-situ measurement. The systematic offset $\overline{\beta^2_{PVA}}$ due to PVA is corrected for the variance measurement based on existing techniques (Frehlich, 1997; Frehlich and Cornman, 2002; Frehlich et al., 2006; Brugger et al., 2016), which are modified for usage with ADL. The correction utilizes an idealized von-Karman spectrum of turbulence, for which the effect of PVA is known given the lidar pulse length and range gate spacing. A random uncertainty $\overline{\varepsilon^2_{PVA}}$ is introduced by the correction method for the second order moment due to a potential violation of the idealized von-Karman spectrum assumed for correction.

(VIII) Finally, due to the limited sample size as well as gaps in the lidar data (e.g due to blocking by clouds or low SNR), a number of data processing procedures have to be applied in order to obtain a meaningful statistical and spectral analysis. Here, I follow the techniques outlined by Kiemle et al. (2011). In this line, for the calculation of statistical

moments, due to the averaging over a finite time series, the desired ensemble mean of a moment is replaced with the temporal mean of the moment. This replacement introduces systematic and random error. The estimation of the associated systematic offset $\overline{\beta^2_{sys}}$ (induced for the variance only) and random uncertainties $\overline{\varepsilon^1_{ran}}$ and $\overline{\varepsilon^2_{ran}}$ is conducted based on Lenschow and Stankov (1986) and Lenschow et al. (1994).

In Chap. 6 it is evaluated to what degree the mentioned sources of errors discussed above are relevant and can be represented using uncertainty estimation in real-world measurements. In order to do so, the specific theories associated with each of the uncertainty estimation methods are outlined in the respective sections. Overall, the results allow for quantification of the uncertainty associated with each source of error for the first-order and second-order moment. Besides this demonstration, the results also allow a discussion of the assumption to neglect the effects applicable for real-world ADL measurements in the ADLS simulations, proving its validity.
For the combined uncertainty evaluation of the mean the concept by Damiani and Haimov (2006) is used and extended. Errors in the motion corrected velocity $\overline{V^1_{COR}}$ can be associated with any of the components in the motion correction equation (Eq. 3.12) and their effect can be systematic or random with either sign. Namely, these are the measured radial Doppler velocity itself, the beam pointing direction and the aircraft velocity projection and thereby also the time synchronization. Further, for vertical wind measurements, a motion and horizontal wind corrected velocity $\overline{V^1_{PST}}$ is required in which the atmospheric contribution due to horizontal wind is removed. This procedure introduces additional uncertainty due to the terminal fall velocity of the scatterers, as well as uncertainty in the knowledge of the wind profile contribution. Last, any measurement of the mean vertical wind has to account for the random uncertainty introduced by substituting the ensemble mean with the temporal mean. Therefore, the following relationship is obtained:

$$\overline{w^1} = \overline{V^1_{PST}} \pm \overline{\beta^1_t} \pm \overline{\beta^1_b} \pm \overline{\beta^1_{rv}} \pm \overline{\beta^1_{ac}} \pm \overline{\beta^1_{v_t}} \pm \overline{\beta^1_{HC}} \pm \overline{\varepsilon^1_t} \pm \overline{\varepsilon^1_b} \pm \overline{\varepsilon^1_{rv}} \pm \overline{\varepsilon^1_{ac}} \pm \overline{\varepsilon^1_{v_t}} \pm \overline{\varepsilon^1_{HC}} \pm \overline{\varepsilon^1_{ran}}. \tag{3.31}$$

For the combined uncertainty evaluation of the variance the concept outlined by Strauss et al. (2015) is followed and extended, but with adapted notation combining Lenschow et al. (1994), Frehlich (1997) and Lenschow et al. (2000). Based on the above discussion, extending Strauss et al. (2015), the measured, motion and horizontal wind corrected velocity variance $\overline{V^2_{PST}}$ can be related to the desired atmospheric vertical wind variance $\overline{w^2}$ through

$$\overline{w^2} = \overline{V^2_{PST}} - \overline{\beta^2_t} - \overline{\beta^2_b} - \overline{\beta^2_{ac}} - \overline{\beta^2_{v_t}} - \overline{\beta^2_{rv}} + \overline{\beta^2_{PVA}} - \overline{\beta^2_{HC}} + \overline{\beta^2_{sys}} \pm \overline{\varepsilon^2_{rv}} \pm \overline{\varepsilon^2_{PVA}} \pm \overline{\varepsilon^2_{HC}} \pm \overline{\varepsilon^2_{sys}} \pm \overline{\varepsilon^2_{ran}}. \tag{3.32}$$

As outlined above, the systematic offsets due to the additional variance contributions $\overline{\beta_{rv}^2}$,$\overline{\beta_{PVA}^2}$ and $\overline{\beta_{sys}^2}$ can be estimated using existing theory. Consequently, these contributions can be removed or added depending on their effect, specifically for every measurement situation. However, the estimation of $\overline{\beta_{rv}^2}$,$\overline{\beta_{PVA}^2}$ and $\overline{\beta_{sys}^2}$ itself is also associated with uncertainties $\overline{\varepsilon_{rv}^2}$,$\overline{\varepsilon_{PVA}^2}$ and $\overline{\varepsilon_{sys}^2}$, which still need to be accounted for (and for which no estimation methods exists in literature to my knowledge). The lower bounds of $\overline{\beta_t^2}$,$\overline{\beta_b^2}$,$\overline{\beta_{ac}^2}$,$\overline{\beta_{v_t}^2}$ and $\overline{\beta_{HC}^2}$ are zero as they can only yield positive contributions, which makes their description as a potential bias adequate. However, their systematic effect cannot be corrected based on existing theory, it can therefore only be estimated with an upper bound. The uncertainties $\overline{\varepsilon_{HC}^2}$ and $\overline{\varepsilon_{ran}^2}$ can present positive or negative contributions, which mandates their description as random uncertainties.

Some of the uncertainty contributions can be constrained further specifically for every measurement situation. Any error in time synchronization, beam pointing-angle calibration or INS-based motion correction results in the measured ground return velocities $V_{COR,RAW}(R_{GRD})$ not being $0\,\mathrm{m\,s^{-1}}$, as would be the case for an ideal system measuring a stationary target (after motion correction). Additional effects such as erroneous ground height identification and associated lidar chirp can also result in a non-zero ground return velocity (Appendix A.5). Hence, the measured ground return velocity presents a conservative estimation of the former three errors (the actual error is being bounded because the ground return velocities present an independent reference measurement). Thereby, in this work the uncertainty contributions by a number of terms are conservatively bounded based on the error determined from the measured ground return velocities after motion correction:

$$|\overline{\beta_t^1} \pm \overline{\beta_b^1} \pm \overline{\beta_{ac}^1} \pm \overline{\varepsilon_t^1} \pm \overline{\varepsilon_b^1} \pm \overline{\varepsilon_{ac}^1}| \leq |\overline{V_{COR,RAW}^1}(R_{GRD})|, \tag{3.33}$$

$$\overline{\beta_t^2} + \overline{\beta_b^2} + \overline{\beta_{ac}^2} \leq \overline{V_{COR,RAW}^2}(R_{GRD}). \tag{3.34}$$

The measured ground return velocities allow for differentiation between various aircraft flight maneuvers, e.g. between nadir and turns. As a result, a flight state dependent assessment of the INS error is obtained. In an extension of the same argument, the measured ground return velocities are then used in a second step to correct the detected errors due to the above terms, which strongly improves the measurement accuracy in a refined motion correction procedure (Sec. 6.2.IV).

In a similar manner, the error in the horizontal wind removal can be bounded conservatively by a comparison of the leg-averaged vertical wind mean to a reference value of $0\,\mathrm{m\,s^{-1}}$. This value is mandated by theory, as the mean vertical wind should disappear when averaged over extended areas for continuity reasons. It can therefore also be treated as a reference truth and used to

obtain a conservative error estimate of the horizontal wind contribution removal accuracy for the first-order moment of the vertical wind (Sec. 6.2.VI). Additionally, this conservative estimation also includes all contributions from the error sources (I-V).

3.6. Wind profiling retrieval error and its representation in a data-driven uncertainty estimation method

Besides the errors introduced by the measurement platform, another important source of error exists for AVAD wind profiling due to the retrieval process. As mentioned above, errors are introduced in the retrieved profile by the inversion process if the assumed homogeneous flow conditions are not present in the retrieval volume (Sec. 2.4). A violation of the flow homogeneity assumption is frequently encountered for measurements inside the BL, where the wind field typically deviates from the mean flow state throughout the retrieval volume due to turbulence. Therefore, in this section the above retrieval formalism (Sec. 3.3) is extended to allow for uncertainty assessment using a covariance matrix based uncertainty estimation.

The wind profile retrieval error is equally present in both ADLS simulated and ADL real-world data, as the LES sufficiently represents the turbulent structures (>O(100 m)) responsible for retrieval error, given the used LES and measurement system setup (Sec. 3.7, 4.1, 4.2). Hence, with the advantage of knowing the exact LES input truth in the ADLS, the retrieval error of ADL systems for wind profiling in inhomogeneous flow can be investigated in Chap. 5. Based on this investigation, the functionality and applicability of a data-driven uncertainty estimation method presented here is evaluated in Sec. 5.4.1 and 5.4.2.

A method to assess the uncertainty in the retrieved wind profiles is through uncertainty propagation using the covariance matrix obtained by the inversion (Leon and Vali, 1998; Leon et al., 2006; Menke, 2012; Wang et al., 2015). The covariance matrix of the model parameters can be calculated as

$$\mathbb{C}_M = \overline{\mu_d^2}(\mathbb{G}'\mathbb{G})^{-1}, \tag{3.35}$$

where $\overline{\mu_d^2}$ is the real data variance (assuming uncorrelated data with equal variance) which cannot be determined from the measurements. Consequently, for uncertainty estimation an estimator of the data variance based on the measurements needs to be specified. Leon and Vali (1998) (for AVAD) and Leon et al. (2006) (for airborne dual-Doppler) suggest to use the LSQ-fit residual velocity variance as a proxy for the unknown data variance. However, the validity of this suggestion has not been checked up to date because it is near-impossible to do so in

real-world measurements. Extending and validating ideas by Leon and Vali (1998) and Leon et al. (2006), this work details under what circumstances the LSQ-fit residual variance can be used as a proxy for the unknown data variance. ADLS results show that in order to achieve a successful uncertainty estimation (representative of the observed error), corrections of the residual variance are necessary due to the spatially neighboring and thereby auto-correlated radial velocity measurements, which have not been discussed in-depth previously. Equation (3.35) can be extended to also take into account measurement system errors, e.g. due to motion correction error or lidar radial velocity noise (see Leon et al. (2006)[Eq. 12] or Wang et al. (2015) for a full derivation including measurement system error). However, this extension is not done here for two reasons: First, the focus of the ADLS work is on evaluating wind profiling error due to atmospheric flow inhomogeneity, which is shown to be a major contributor in Chap. 5. Second, the results from Chap. 6 show that compared to the atmospheric retrieval error contribution, typical measurement system errors are small and can be reliably controlled.
Once an estimate of the covariance matrix $\mathbb{C}_M$ is obtained, the uncertainties in the retrieved wind components are then given by the diagonal entries of the covariance matrix (Menke, 2012):

$$\varepsilon_u = \sqrt{\mathbb{C}_M(1,1)} \text{ and } \varepsilon_v = \sqrt{\mathbb{C}_M(2,2)} \text{ and } \varepsilon_w = \sqrt{\mathbb{C}_M(3,3)}. \tag{3.36}$$

The obtained uncertainties can be propagated into the wind speed using (Wang et al., 2015; Newsom et al., 2017):

$$\varepsilon_{V_m} = \sqrt{(u\varepsilon_u)^2 + (v\varepsilon_v)^2 + (w\varepsilon_w)^2}/V_m, \tag{3.37}$$

if no correlation between the u, v and w wind components is assumed and the off-diagonal terms are neglected. Similarly, for the wind direction the propagation is conducted according to (Newsom et al., 2017):

$$\varepsilon_{V_\alpha} = \sqrt{(u\varepsilon_v)^2 + (v\varepsilon_u)^2}/V_m^2. \tag{3.38}$$

In this work, the scale of the flow inhomogeneity causing wind profiling retrieval error is set in proportion to the scale of the measurement volume, similar to the approaches used by Koscielny (1984), Eberhard et al. (1989), Boccippio (1995), and Tian et al. (2015). Thereby, so-called sub-scan-volume flow inhomogeneity is distinguished from so-called supra-scan-volume flow inhomogeneity for reasons provided in the following discussion.

Resolved sub-scan-volume flow inhomogeneity contribution

Flow inhomogeneity with characteristic scales smaller than the scan volume (termed sub-scan-volume or resolved in the following) can cause profiling error. Sub-scan-volume flow inhomogeneity is detectable through the deviations of the measured radial velocities from the LSQ-fit (Fig. 3.4). An estimate of the wind profile uncertainty due to flow inhomogeneity is available through covariance matrix uncertainty propagation (Eq. 3.35), if the deviations can be used as a proxy for the real, unknown data variance. Therefore, for this estimation the unknown data variance is frequently replaced with the residual velocity variance which can be obtained from the LSQ-fit (Leon and Vali, 1998; Leon et al., 2006; Menke, 2012; Wang et al., 2015):

$$\overline{\mu_d^2} = \frac{1}{N-m} \sum_{n=1}^{N} \left(v_{COR_n} - v_{COR_n}^R \right)^2 . \tag{3.39}$$

The above simplified version assumes independent data samples, a condition which is violated for high frequency Doppler lidar radial velocity measurements often conducted in close spatial and temporal proximity to each other. When the separation of measurements is smaller than the temporal or spatial integral length scales of turbulence (quantifying when all memory is lost) the samples are not statistically independent. In particular, any flow inhomogeneity will preserve over the course of many adjacent measurements, leading to structured deviations of the measurements around the LSQ-fit (see Fig. 4.4 for an illustration and Eberhard et al. (1989) for a similar approach). This lack of statistical independence results in strong (auto-) correlation in the residuals. Thereby, the radial velocity measurements from one or more range gates do not necessarily add new, independent information. As a result, the effective sample size is strongly reduced, an effect which needs to be accounted for in the uncertainty estimation.
For along-beam neighboring range gates (sampled at the same time), the reduced information content is accounted for by reducing the sample size N depending on the degree of correlation in the residuals from neighboring range gates, termed ρ_{res} here:

$$N_{corr} = N/[1 + (N-1)\rho_{res}]. \tag{3.40}$$

For measurements from the same range gate, conducted after each other in close spatial and temporal proximity, the correlation of the measurements is accounted for by replacing N_{corr} with an effective sample size N_{eff}, which reflects the reduced number of samples depending on

the amount of correlation present. In this work, the estimator used for the effective sample size is based on Thiébaux and Zwiers (1984) and Storch and Zwiers (1999):

$$N_{eff} = N_{corr}/[1 + 2\sum_{\tau=1}^{N-1}(1 - \frac{\tau}{N}\rho_{aco}(\tau))]. \tag{3.41}$$

Therein, $\rho_{aco}(\tau)$ is the autocorrelation function of the residuals at lag τ. The autocorrelation cannot be calculated directly from the overall autocorrelation function of the residuals of the LSQ-fit. The problem is that multiple range gates can fall into one retrieval volume, with the different range gates exhibiting different residual behavior (as they sample different turbulent structures). Further, for an AVAD technique, the range gates can sample the same retrieval volume multiple times at different times (e.g. for an forward and aft beam). Therefore, the autocorrelation function is calculated as a weighted average from all range gates in the retrieval volume. To do so, individual autocorrelation functions $\rho_{aco}(\tau, n_R)$ are calculated individually for the number of range gates n_R falling into one volume. Measurements from the same range gate but separated by more than 2 s are treated individually. Subsequently, the individual autocorrelation functions are weighted by the number of their contributing measurements $N(n_R)$ compared to the overall number of measurements and summed up:

$$\rho_{aco}(\tau) = \sum_{n_R}(\rho_{aco}(\tau, n_R) \cdot N(n_R)/N)\,. \tag{3.42}$$

The uncertainty estimation method presented here is expected to represent errors due to sub-scan-volume flow inhomogeneity. The method can be challenged for three reasons. First, it remains to be checked to what extent the residual velocity variance is representative of the underlying, unknown data variance. Second, the assumption of a single, average autocorrelation function representing the reduction in sample size is a simplification. As stated above, in reality the degree of autocorrelation is not constant but depends on the structure of turbulence as well as the timing and location of measurements. Further, the covariance matrix is not necessarily diagonal in reality but may contain off-diagonal terms as well. Third, estimating the effective sample size from data is not a trivial task and multiple approaches and formulas exist (Thiébaux and Zwiers, 1984). Generally, estimation is more challenging for small sample sizes (N<30) and depends on the stochastic nature of the process observed. The approach chosen here is a simple version and more sophisticated approaches could be included in future work. To gauge the effect and magnitude of these problems a validation of the uncertainty estimation method applicability is conducted using the ADLS in Sec. 5.4.2.

Generally, as mentioned above, only flow inhomogeneity smaller than the size of the scan volume is detectable as residuals in the LSQ-fit. Therefore, any errors due to large scale flow inhomogeneity cannot be captured by the covariance uncertainty propagation and need to be evaluated separately. Methods to do so in the ADLS and in real-world measurements are presented in the next section.

Unresolved supra-scan-volume flow inhomogeneity contribution

Flow inhomogeneity on the order of the scan volume and beyond (termed supra-scan-volume in the following) is also a source of wind profiling error (Koscielny, 1984; Boccippio, 1995; Tian et al., 2015). An illustrative example is Doppler lidar wind profiling in a horizontally sheared vertical wind field with $0\,\mathrm{m\,s^{-1}}$ horizontal background wind (Fig. 3.4). The radial velocity contribution of the vertical wind is erroneously mapped into horizontal wind (depending on the vertical wind magnitude and measurement geometry, e.g. the scan elevation angle), if a VAD, VVP or AVAD retrieval algorithm is applied. Hence, the supra-scan-volume contribution is also called erroneous mapping in the following. Erroneous mapping can also occur under the presence of background wind. An illustrative example is a case where the vertical wind is in-phase with the horizontal wind contribution. For example, this alignment can occur in situations when there is an updraft upstream of the lidar and a downdraft downstream (again a version of the horizontally sheared vertical wind field, the overall vertical wind is zero). In this case, the vertical wind projection appears with the same azimuthal signature as the horizontal wind in the measured radial velocities used in the AVAD. Consequently, the vertical wind is mapped into a horizontal wind contribution, causing a retrieval error.

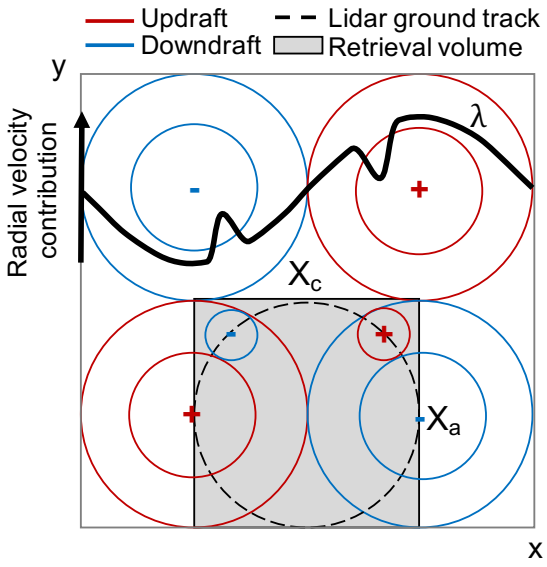

Fig. 3.4.: Schematic depiction of a Doppler lidar VAD wind profile retrieval being impacted by flow inhomogeneity. The supra-scan-volume inhomogeneity of the vertical wind is erroneously mapped into horizontal wind because it is exactly in-phase with the measurement geometry. The sub-scan-volume flow inhomogeneity causes detectable residuals in a LSQ-fit. For simplicity a non-moving lidar system is shown.

Supra-scan-volume contributions to wind profiling error do not cause detectable residuals of the fitted radial velocity from the measured radial velocity (as they are unresolved and erroneously mapped). The wind profiling error is thereby not detectable through the covariance matrix uncertainty propagation, as is done for sub-scan-volume flow inhomogeneity. For individual wind profiles the erroneous mapping due to supra-scan-volume flow inhomogeneity can only be detected and quantified using the ADLS because the three-dimensional wind along the measurement trajectory is needed, which is nearly impossible to obtain in real-world lidar measurements. In the simulation results, cases with almost exact phase-alignment (between vertical wind and scan geometry) can be observed, Consequently, these LSQ-fits show very high R^2 values with little estimated sub-scan-volume uncertainty, but nevertheless exhibit a large wind profile error ($> 1\,\mathrm{m\,s^{-1}}$).

Besides the strength and extent of the flow inhomogeneity, the degree and process of erroneous mapping depends on the measurement geometry used (Koscielny, 1984; Boccippio, 1995). At constant elevation, only changes in the vertical wind cause erroneous mapping. However, if the beam elevation is varied during the scan, changes in the horizontal wind components can also cause erroneous mapping. Erroneous mapping is more prominent at high elevation angles when the vertical wind exhibits a large contribution to the measured radial velocity compared to the horizontal wind. It can occur even for large measurement volumes (long sampling duration, low spatial resolution), for example when flying perpendicular to the wind with presence of organized convective structures or in complex terrain.

Because erroneous mapping by the retrieval is not captured by the covariance matrix based residual variance propagation, two methods to account for wind profiling uncertainty due to supra-scan-volume flow inhomogeneity are proposed. The first method is only applicable when the atmospheric input data is known exactly (e.g. in the ADLS). The second method parameterizes the uncertainty due to supra-scan-volume flow inhomogeneity in order to represent it on a statistical basis in real-world BL measurements. Thereby, a data-driven uncertainty estimation of wind profiling error can be achieved. The mathematical details on the quantification of the supra-scan-volume contributions and its parameterization are provided in Appendix A.3.

In practice the separation between the sub- and supra-scan-volume contributions is not always clear and easy to make. The scanning geometry can be complex and the turbulent spectrum is continuous and does not exhibit the same separation. Therefore, the developed theory needs to be checked and tested. The ADLS provides an ideal testbed to do so as the input truth is known exactly, which is near-impossible in real-world measurements. Hence, the validity of the proposed uncertainty estimation method can be evaluated (Sec. 5.4.2).

3.7. Large Eddy Simulations

In order to evaluate measurement system characteristics and obtain a better handle of retrieval errors a simulation based measurement system analysis is useful (Sec. 2.4.2). In order to do so, artificially generated high resolution atmospheric wind fields are necessary. To capture the effect of turbulence on ADL measurements the wind fields must have sufficient resolution to represent turbulence on the scale of the lidar measurement geometry. At the same time, available computational power must allow for simulation of a large enough domain for statistical analysis of ADL wind profiling retrieval.

The flow is described by the NVS equations (Eq. 2.1), which allow for analytical solution only with application of simplifying assumptions. Thus, the equations are solved numerically for the simulation of turbulent wind fields.

Turbulence resolving, direct numerical simulation (DNS) of fluids through solving the full set of equations is numerically expensive. Only small domains (O(10 m)) can be simulated using DNS, due to the fine resolution needed in order to resolve the full spectrum down to the smallest turbulent elements (O(1 mm), besides other processes such as cloud micro-physics on even smaller scales). Therefore, DNS is not appropriate to generate the input wind fields needed in this work.

Besides DNS, a variety of numerical models exist which focus on simulating atmospheric flow on larger scales. The different models introduce approximations into the NVS equations based on scale analysis and thereby allow for simplified simulation. For example, models used for NWP focus on synoptic scales using the so-called Reynolds-Averaged Navier-Stokes Equations (RANS, Sec. 2.1). In RANS the scale separation is in the region of the atmospheric spectral gap, distinguishing mean flow from turbulent flow. Current typical NWP grid spacing is O(10 km) globally (Deutscher Wetterdienst, 2020a) and O(1 km) locally, using so-called nesting approaches (Gasch et al., 2017; Klocke et al., 2017). As a resulting drawback of the coarse grid spacing, NWP models cannot resolve atmospheric BL processes such as turbulence (their effect is included through parametrizations).

Situated in the range between DNS and NWP models, the large eddy simulation approach exists with a grid spacing O(1 – 100 m). For LES models the separation between resolved and unresolved scales is located in the inertial subrange part of the turbulent spectrum. Thereby, the important production range of the turbulent spectrum is grid-scale (GS) and can be resolved over larger domains and at computational cost lower than for DNS. Turbulence at smallest scales, e.g. the lowest part of the inertial subrange and the dissipation range, need to be parameterized using a subgrid-scale (SGS) model in LES. The capability of the LES to resolve the produc-

tion range and most of the inertial subrange allows for simulation of the relevant part of the turbulent spectrum driving wind profiling retrieval error. Further, today's computational power is sufficient to enable generation of sufficiently large wind fields for statistical analysis of wind profiling. The parameterization of turbulence on the smallest scales does not prevent a simulation of the Doppler lidar measurements, as it occurs on scales finer than the range resolution of the Doppler lidar itself in the chosen setup (Sec. 4.1).

3.7.1. The PALM model

The Parallelized Large-Eddy Simulation Model (PALM) was first developed at the Leibniz University Hannover by Raasch and Etling (1991). Since then it has undergone numerous developments, e.g. parallelization (Raasch and Schröter, 2001) and inclusion of other packages such as a Lagrangian particle and cloud model, a canopy model and oceanic turbulence. An overview of available features and studies utilizing PALM is available in Maronga et al. (2015, 2020) and the following description is based on the description provided there.
The equations solved by PALM for the GS variables are derived from the NVS equations (Eq. 2.1) utilizing the Boussinesq approximation (Etling, 2008; Maronga et al., 2020). In the following equations, summation over repeated indices is implied, overlines indicate GS variables, primes indicate SGS variables, the subscript 0 indicates surface variables, angle brackets denote a horizontal domain average. For a dry atmosphere (humidity is not relevant for this work) the equations read as:

$$\frac{\partial \overline{u}_i}{\partial t} = -\frac{\partial \overline{u}_i \overline{u}_j}{\partial x_j} - \varepsilon_{ijk} f_i \overline{u}_k + \varepsilon_{i3j} f_3 U_{g,j} - \frac{1}{\rho_0} \frac{\partial \overline{p^* + \frac{2}{3}\rho_0 E_{SGS}}}{\partial x_i} + g \frac{\overline{\Theta} - \langle \overline{\Theta} \rangle}{\langle \overline{\Theta} \rangle} \delta_{i3} - \frac{\partial}{\partial x_j} \left(\overline{u_i' u_j'} - \frac{2}{3} E_{SGS} \delta_{ij} \right), \tag{3.43}$$

$$\frac{\partial \overline{u}_i}{\partial x_j} = 0, \tag{3.44}$$

$$\frac{\partial \overline{\Theta}}{\partial t} = -\frac{\partial \overline{u}_j \overline{\Theta}}{\partial x_j} - \frac{\partial}{\partial x_j} \left(\overline{u_j' \Theta'} \right) - Q_\Theta. \tag{3.45}$$

Additionally, for turbulence closure a prognostic equation for the SGS-TKE is formulated as

$$\frac{\partial E_{SGS}}{\partial t} = -\overline{u}_j \frac{\partial E_{SGS}}{\partial x_j} - \left(\overline{u_i' u_j'} \right) \frac{\partial \overline{u}_i}{\partial x_j} + \frac{g}{\Theta_0} \left(\overline{u_3' \Theta'} \right) - \frac{\partial}{\partial x_j} \left[\overline{u_j' \left(E_{SGS} + \frac{p'}{\rho_0} \right)} \right] - \epsilon. \tag{3.46}$$

Tab. 3.1.: Overview of surface and top boundary conditions applied in both LES sets A and B.

	Surface boundary	Top boundary
Horizontal wind	$u = v = 0\,\mathrm{m\,s^{-1}}$	$u = u_G; v = v_G$
Vertical wind	$w = 0\,\mathrm{m\,s^{-1}}$	$w = 0\,\mathrm{m\,s^{-1}}$
Potential temperature	$\frac{\partial \Theta}{\partial z} = 0\,\mathrm{K\,m^{-1}}$	$\frac{\partial \Theta}{\partial z} = \frac{\partial \Theta}{\partial z}_{initial}$
TKE	$\frac{\partial E_{SGS}}{\partial z} = 0\,\mathrm{m\,s^{-2}}$	$\frac{\partial E_{SGS}}{\partial z} = 0\,\mathrm{m\,s^{-2}}$
Pressure	$\frac{\partial p}{\partial z} = 0\,\mathrm{hPa\,m^{-1}}$	$\frac{\partial p}{\partial z} = 0\,\mathrm{hPa\,m^{-1}}$

The equations contain five prognostic variables, these are the three wind components u_i with $i = 1, 2, 3$, the potential temperature Θ and the SGS turbulent kinetic energy $E_{SGS} = \frac{1}{2}\overline{u_i'^2}$ (six variables if the specific humidity q_v or another passive scalar is included). Compared to Eq. 2.1, new variables are the perturbation pressure p^*, the density of dry air ρ_0, the source term Q_Θ and the SGS dissipation rate ϵ. Corresponding, $U_{g,j}$ are the components of the geostrophic wind.

The model uses an equidistant staggered Arakawa C-grid (scalar quantities at cell center, velocity components shifted by half a grid spacing) for discretization of the fields. The advection terms are discretized using a fifth-order differencing scheme in space and a third-order Runge-Kutta scheme in time.

As discussed, an important part of the LES is the parameterization of SGS turbulence (a RANS version of PALM with fully parameterized turbulence is also available). The PALM model formulation contains SGS covariance terms which need to be parameterized (for the SGS turbulent transport of momentum, heat and potentially humidity or another passive scalar, further the pressure term in the SGS-TKE equation) as well as the dissipation ϵ. In PALM, turbulence closure is achieved by using a one-and-a-half order closure based on Deardorff (1980). The contribution by the non-resolved terms is included using a gradient based approach, which assumes that the energy transport by SGS turbulence is proportional to the local gradient of the mean GS quantities. For the pressure term the transport is assumed proportional to the gradient of SGS-TKE and the dissipation is set proportional to $E_{SGS}^{\frac{3}{2}}$ (and a factor taking into account the grid spacing).

The Boussinesq approximated equations require incompressibility, which is not achieved by the integration per se. Therefore, a predictor-corrector method is applied in which the pressure term is excluded during the first integration of the NVS equations to yield a preliminary solution for u_i. Developing divergences are then attributable to the pressure term. Reinforcing incompressibility through stipulation results in a Poisson equation for the pressure which is then solved using an FFT procedure.

In order to conduct simulations boundary conditions must be specified. The simulations used in this work assume a flat, homogeneous surface and a constant heating rate/heat flux for LES set A/set B, respectively. An overview of the boundary conditions applied at the surface and top is provided in Table 3.1. Laterally, periodic boundary conditions are used in both directions, e.g. the outflow from the domain is used as inflow again for all variables. The model state at initialization is prescribed through vertical profiles for all variables (horizontally homogeneous). To initiate the development of turbulence small random perturbations are introduced into the wind fields until a steady turbulence develops.

4. An LES-based airborne Doppler lidar simulator [1]

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 ADL 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. A sketch of the ADLS operating scheme used for wind profiling is provided in Fig. C.1, it is detailed and explained in the following.

4.1. LES generated wind fields

In order to be as close as possible to a realistic measurement environment the work presented here utilizes high resolution LES generated wind fields. The LES wind fields are simulated using PALM and provided by the University of Hanover (Raasch and Schröter, 2001). 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. On the other hand, the domain has to be large enough to ensure a sufficient number of independently sampled wind profiles for statistical analysis.

Two different sets of LES simulations are used in this work (set A, set B). Both sets employ a range of background wind speeds and a grid spacing of $\Delta x = 10\,\mathrm{m}$ (corresponding to a resolution finer than O(100 m)). The main difference between LES set A and set B is the surface sensible heat flux, which is approx. $170\,\mathrm{W\,m^{-2}}$ for set A and $30\,\mathrm{W\,m^{-2}}$ for set B, leading to different vertical BL and turbulence structures.

LES wind fields resolve turbulence appropriately when the resolved grid-scale turbulent kinetic energy E_{GS} is much larger than parameterized subgrid-scale turbulent kinetic energy E_{SGS}. Given the LES setup and turbulence present, this condition is fulfilled for all used LES wind fields above an altitude of 30 m (Fig. C.2). Further, at the assumed lidar range gate length $\Delta p =$

[1] The chapter is adapted from the publication: Gasch, P., Wieser, A., Lundquist, J. K. and Kalthoff, N., 2020: An LES-based airborne Doppler lidar simulator and its application to wind profiling in inhomogeneous flow conditions. Atmospheric Measurement Techniques, 13(3), 1609-1631.

Tab. 4.1.: Overview of atmospheric conditions present in LES sets A and B. The BL height is determined from the potential temperature profile.

	LES set A					LES set B		
Grid spacing in m	10					10		
Domain size in m^3	23030 x 17270 x 2300					5000 x 5000 x 1800		
Simulation duration in min	120					30		
Output temporal resolution in s	60					1		
Background wind speed in $m\,s^{-1}$	0	2	5	10	15	5	10	15
$\overline{w'\Theta'}$ in $K\,m\,s^{-1}$	0.13	0.15	0.16	0.17	0.17	0.03	0.03	0.03
u_* in $m\,s^{-1}$	0.17	0.22	0.37	0.62	0.85	0.32	0.51	0.68
w_* in $m\,s^{-1}$	1.68	1.74	1.76	1.83	1.90	0.84	0.85	0.87
$-z_i/L_0$	275	157	37	9.6	4.3	7.2	1.9	0.8
z_i (Θ) in m	1100	1100	1120	1200	1270	1142	1114	1152

69.53 m (Sec. 3.1), the simulations also fulfill the condition $\Delta p >> \Delta x$ and thereby Doppler lidar measurements can be represented realistically (Stawiarski, 2014).

The boundary conditions applied in both sets at the surface and top are summarized in Table 3.1. Laterally, periodic boundary conditions are used in both directions, e.g. the outflow from the domain is used as inflow again for all variables. Although the domain size and simulation time vary, both LES set A and set B allow for sampling 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. 4.4.2.

The ADLS can be easily adopted for use with other LES wind fields (e.g. different atmospheric conditions, larger or longer simulation domains), should this be necessary for future studies.

LES set A

The LES set A is provided by Oliver Maas and Prof. Siegfried Raasch, PALM group, Institute of Meteorology and Climatology at Leibniz Universität Hannover. The simulations are conducted using PALM version 6.0.

The LES set A employs a large domain size of 23030 m x 17270 m x 2300 m. Vertically, the output is limited to 1500 m. Vertical profiles of the average wind speed, the potential temperature, the kinematic sensible heat flux and the component-wise wind variance are provided in Fig. C.2. An overview of other relevant meteorological parameters is provided in Table 4.1. LES set A is driven with five geostrophic background wind speeds of $u_G = 0, 2, 5, 10$ and 15 $m\,s^{-1}$ and simulates a dry atmosphere. At the flat surface a constant heating rate of $+0.6\,K\,h^{-1}$ is prescribed, which results in a kinematic heat flux of approx. $0.16\,K\,m\,s^{-1}$ (corresponding to a sensible heat flux of approx. $160\,W\,m^{-2}$, with some variation between the background wind

cases). In order to dampen the inertial oscillation of the geostrophic wind, a 12 h pre-run is conducted with a reduced simulation domain. After a subsequent spin-up time of 4 h using the full domain (leading to the decay of periodic structures present in the LES), three-dimensional data output started with fully developed turbulence at a temporal resolution of 1 min. In total, 2 h of data output are available. The convective situation is classified as strongly unstable stratification with the stability criteria pointing to the development of organized convective structures for $u_G > 0\ \mathrm{m\,s^{-1}}$. The BL height is between 1100 – 1300 m, with the entrainment zone extending from 1100 – 1400 m. Gravity waves are present in the entrainment zone and above.

LES set B

The LES set B is provided by Dr. Christoph Knigge and Carolin Helmke, PALM group, Institute of Meteorology and Climatology at Leibniz Universität Hannover. The simulations are conducted using PALM version 3.9.
The LES set B used in this work has a smaller domain size of 5000 m x 5000 m x 1800 m. The set is 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) including the vertical profiles (their Fig. 1). An overview of other relevant meteorological parameters is provided in Table 4.1. Set B also simulates a dry atmosphere, it is driven with three geostrophic background wind speeds of $u_G = 5, 10$ and $15\ \mathrm{m\,s^{-1}}$ and a constant kinematic surface heat flux $\overline{w'\Theta'} = 0.03\ \mathrm{K\,m\,s^{-1}}$ at the flat surface. Data output started with fully developed turbulence at a temporal resolution of 1 s after a spin-up time of 1 h. The convective situation is classified as unstable stratification with the stability criteria pointing to the development of organized convective structures. The kinematic heat flux profile approaches zero at a height of approximately 600 m. An entrainment zone with an associated slightly positive potential temperature gradient extends from 600 – 1100 m. Both turbulent structures and gravity waves are present in the entrainment zone and above, the BL height as determined from the potential temperature profiles is approximately 1100 m (Stawiarski, 2014).

4.2. System - idealized airborne Doppler lidar platform

In this section the geometric simulation of the main system components consisting of aircraft, scanner and lidar is discussed. An illustrative overview of the results obtained after system simulation is shown in Fig. 4.1. A so-called 'ideal' system is presented for usage in the ADLS in the following. 'Ideal' as used here refers to the absence of any errors in the state of the system components (measurement platform errors, Sec. 3.5), as their status is completely known (pre-

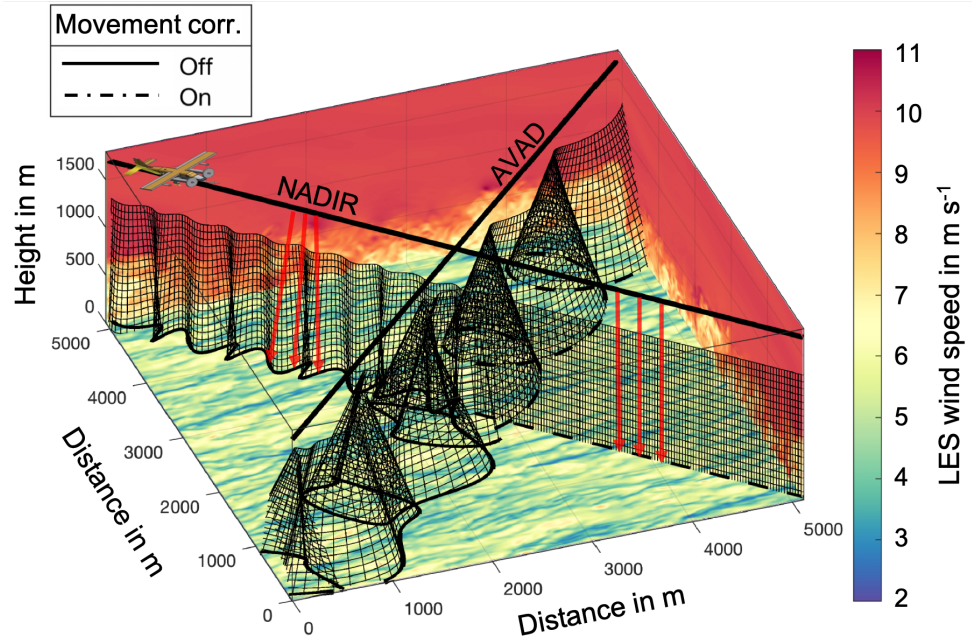

Fig. 4.1.: Illustration of system and measurement position simulation for two transects through the LES domain (for LES set B). The LES wind speed is color-coded. The black lines represent aircraft trajectories. The black curtains show the range gate positions used to conduct the measurement which are 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.

scribed) in the ADLS. Thereby, the evaluation can be focused on the retrieval errors introduced by flow inhomogeneity (Sec. 3.6). While this assumption is not generally valid in real-world ADL measurements, the results presented in Chap. 6 show that the errors introduced by the system components are small and can be reliably evaluated using uncertainty estimation methods. Thus, system errors can neglected compared to the retrieval errors introduced due to flow inhomogeneity, as revealed by the ADLS for turbulent conditions.

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. Another parameter that can be varied is the aircraft flight altitude.
Depending on whether a frozen-in-time (LES set A) or time varying (LES set B) wind field is present during sampling, the aircraft trajectory and sampling positions inside the LES must be calculated differently. The aircraft moves relative to the air mass and not 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 ground or the air mass, different sampling distances have to be applied in the LES. To my 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 during sampling (LES set A), the LES coordinate system coincides with that of the air mass. This approach is often utilized for LES-based investigation of airborne measurement characteristics in literature (Schröter et al., 2000; Sühring and Raasch, 2013; Petty, 2020), due to its simplicity in sampling implementation and because LES output is often not available at sufficiently high temporal resolution. 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 using the aircraft speed relative to air (IAS) through the simple relationship $s = IAS \cdot t$. Consequently, for a given sampling time and IAS 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 A.4), because the aircraft track with respect to the ground is influenced by the wind speed. To illustrate the concept, albeit being unrealistic, consider an aircraft flying at $65\,\mathrm{m\,s^{-1}}$, either up- or downwind, with a LES wind speed of $65\,\mathrm{m\,s^{-1}}$. 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 a time-varying wind field during sampling (LES set B), the aircraft also moves relative to the air mass but now the LES coordinate system coincides with the ground coordinate system. This case is deemed slightly more realistic, as it corresponds directly to real-world measurements situations. In this case, the motion of the aircraft due to the air mass motion has to be considered during each time step to yield the correct measurement positions in the LES. Therefore, 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 LES coordinate system coincides with the ground, the sampling distances applied in the LES have to be adjusted for the actual distance covered by the aircraft with respect to the ground, 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 during each time step to yield the actual measurement positions in the LES. To this end, the aircraft heading and ground speed are determined iteratively for subsequent time steps, based on the triangle of velocities (Appendix A.4). The effect of this procedure can again

be imagined with an aircraft flying at $65\,\mathrm{m\,s^{-1}}$ into $65\,\mathrm{m\,s^{-1}}$ 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.
The correct simulation of the aircraft motion is important as the sampling of the wind field by the lidar system is altered. In this work, wind speeds reach up to 23% of the IAS. Thereby, the sampling distance between measurements is changed by a factor of up to 0.4 between flying up- or downwind ($50\,\mathrm{m\,s^{-1}}$ vs. $80\,\mathrm{m\,s^{-1}}$ ground speed), presenting a non-negligible effect. Additionally to the aircraft track development as described above, pitch, yaw and roll moments can be added to the aircraft position to simulate the effect of aircraft oscillations on the measurement process. These are 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.
In the ADLS an ideal aircraft system is assumed. This means that the aircraft position, orientation and velocity is known precisely and without error at all times (e.g a perfect INS is available). The ideal aircraft system assumption is justified by the real-world ADL results obtained in Chap. 6. The results presented there show that the real INS available can suffer from non-negligible error, especially during turns. However, it is demonstrated that these errors can be further controlled and reduced through usage of a refined motion correction procedure, which utilizes the measured ground return velocities.

Scanner

In anticipation of an upcoming system the scanner movement is simulated with freely selectable scanning geometries and includes an option to correct for the aircraft movement. Although the intended scanning system was not built in-time for inclusion in this work by the manufacturer, due to unforeseen manufacturing and aircraft certification issues, the results and developed techniques are transferable to a fixed-beam system as is investigated in Chap. 6.
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.
In the ADLS an ideal scanner system without beam pointing error is assumed. The ideal scanner system assumption is justified by the real-world ADL results obtained in Chap. 6. There, it is shown that the uncertainty due to beam pointing error can be reliably assessed and controlled using the lidar measured ground return velocities.

Lidar

The simulated lidar is based on the specifications of a Lockheed Martin Coherent Technologies WTX WindTracer Doppler lidar which is also used for real-world measurements in Chap. 6. Lidar systems with similar characteristics are often used in ADL studies (Weissmann et al., 2005b; 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 using a Gaussian pulse width of $\sigma_p = 1.8 \times 10^{-7}$ s is applied (corresponding to a temporal FWHM of $\Delta t = 3 \times 10^{-7}$ s and a spatial FWHM of $\Delta r = 45$ m, see Sec. 3.1). 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 $\Delta p = 69.53$ m with a non-overlapping spacing of $\Delta R_0 = 69.53$ m (the distance between subsequent range gate centers), starting at a distance of 400 m from the lidar is used for this work. The lidar measurement frequency 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.
In the ADLS, an ideal laser system is assumed without random radial velocity fluctuations or other system errors. The ideal lidar assumption is justified by the real-world ADL results obtained in Chap. 6, where it is shown that the random radial velocity variance of the WTX can be reliably assessed and is low under all measurement conditions.

4.3. Measurement procedure

For the measurement, the aircraft and atmospheric motion vectors need to be combined to give the measured radial Doppler velocities using Eq. 3.11, explained in Sec. 3.2. Additionally, to obtain the atmospheric motion vector from the LES, the locations of the lidar range gates need to be calculated. 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.

Figure 4.1 illustrates the measurement locations obtained from this procedure for both nadir and AVAD transects.

In order to determine the LES wind velocity v_p^E which is projected onto the beam, the range gate position has to be determined. 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 (Appendix A.2), the range gate center positions are calculated by adding the range gate center distances R_0 (spaced $\Delta R_0 = \Delta p = 69.53\,\mathrm{m}$ apart) in beam orientation to the aircraft position,

$$\boldsymbol{p}_{R_0}^E = \boldsymbol{p}_{AC}^E + R_0 \boldsymbol{b}^E. \tag{4.1}$$

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

$$R_{0B} = R_0 - \Delta p/2, \tag{4.2}$$

$$R_{0E} = R_0 + \Delta p/2. \tag{4.3}$$

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 needs 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 start and the completion 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 positions compared to the range gate positions of the previous measurement (this assumption is valid as only continuous aircraft and scanner movements are investigated). Using these positions, three individual difference vectors (the range gate motion vectors) are constructed for the range gate begin, center and end position each. The first and last range gate positions during one measurement are then obtained by subtracting and adding half the range gate motion vector from the range gate center, begin and end positions, respectively.

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 (the grid spacing is $\Delta x = 10$ 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 maximum three-dimensional grid point distance is employed to ensure points inside the volume,

$$d_{\text{inflate}} = \sqrt{3}/2\Delta x. \tag{4.4}$$

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 based on their orthogonal distance from the beam center in along-beam direction (in accordance with Stawiarski (2014), see Sec. 3.1). 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. 3.11.
Because measurement platform errors are neglected for the ideal ADLS system, the motion-corrected Doppler velocity due to the particle velocity is equal to the particle velocity projection itself. 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, all wind profiling errors discussed in the next sections stem purely from the inhomogeneous atmospheric flow conditions due to BL turbulence. In a homogeneous wind field, the retrieved wind profile is exact.

4.4. Retrieval

After the simulated measurements are obtained from the ADLS the retrieval can be performed. In this section a nadir retrieval is performed, which illustrates a number of features in the ADLS. Further, examples of ADLS retrieved wind profiles are also shown and the system setup, retrieval strategy and sampling procedure used for the statistical analysis of wind profiling error in Chap. 5 is discussed alongside.

4.4.1. Nadir as an example application

An illustrative example of the ADLS capability is given by simulated nadir retrievals. In Fig. 4.2 the simulated vertical wind measurement is compared to the LES set B input along two transects for an ideal system with a lidar measurement frequency of 10 Hz. The aircraft is flying at an altitude of 1700 m with an IAS of 65 $\mathrm{m\,s^{-1}}$ and a superimposed 8 s, 5° roll oscillation, combined with a 15 s, 3° pitch oscillation.
The first transect, a crosswind flight (Fig. 4.2a, b), reveals the spanwise turbulence structure for the 10 $\mathrm{m\,s^{-1}}$ background wind case (LES set B). Along-stream organization of turbulence into rolls occurs and gravity waves are present above the BL. The ADLS results show that the

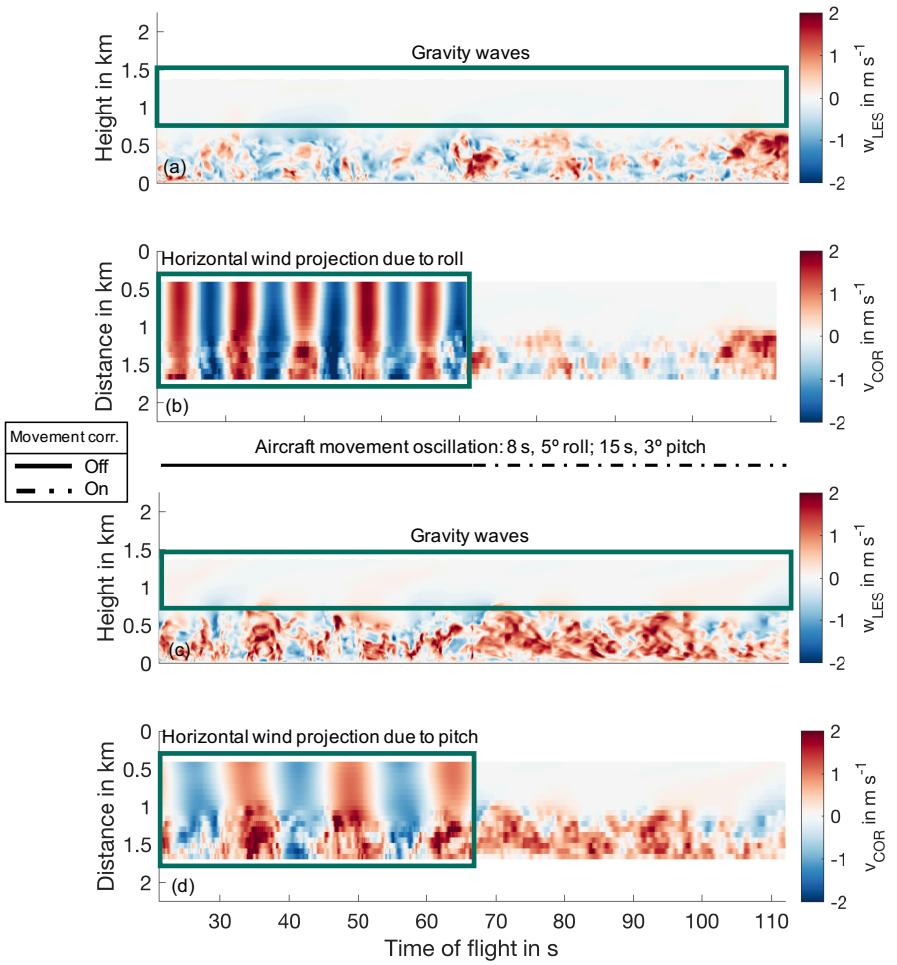

Fig. 4.2.: ADLS simulated nadir transects for the $10\,\mathrm{m\,s^{-1}}$ background wind case of LES set B. The scanner movement correction is disabled for the first half of the transects, whereas it is enabled for the second half. (a) LES vertical wind along a crosswind transect. (b) Corresponding ADLS motion corrected velocity measurement along the crosswind transect. (c) LES vertical wind along an upwind transect. (d) Corresponding ADLS motion corrected velocity measurement along the upwind transect.

simulated lidar is able to capture turbulent structures in the BL, although smoothing occurs for the finest scales. Vertically, the resolution of the measurement is defined by the pulse and 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 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 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 of the horizontal wind can be removed, as a second step, if the wind profile is known (Chouza et al., 2016b). However, deviations of the horizontal wind from the mean 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. 4.1). 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. 4.2c, d), reveals the streamwise turbulence structure for the $10\,\mathrm{m\,s^{-1}}$ background wind case (LES set B). The flight is conducted in an updraft region of the along-stream organized BL convection, therefore positive vertical winds dominate. In this case, the superimposed pitch movement contaminates the vertical wind measurement with a contribution from the horizontal wind if the scanner movement correction is disabled.

Overall, nadir measurements provide a good opportunity to check the accuracy of the scanner movement compensation and beam pointing-angle accuracy using the motion corrected wind measurement (and ground return velocity in real-world measurements), especially if superimposed with controlled pitch and roll maneuvers. If an aircraft movement compensation through the scanner is not available (e.g. for a fixed beam system as investigated in Chap. 6), or if the scanner movement correction is not sufficiently accurate, the amount of horizontal wind mapped into the measurement can be removed in a further post-processing step, if the vertical profile of the horizontal wind is known. In order for this technique to yield reliable results, an accurate wind profile retrieval is necessary. Consequently, the reliability of wind profiling measurements is the focus of the next sections. Results from Chap. 6 also show that a number of other uncertainties present in nadir vertical wind measurements are reliably quantifiable, further motivating the focus on wind profile retrieval error in the following.

4.4.2. Wind profiling

The assessment of wind profiling error in inhomogeneous flow conditions, as well as system setup and retrieval strategy optimization, is a main focus of the ADLS study. In order to conduct this assessment the ADLS is used to simulate a sufficient number of wind profile retrievals for statistical analysis and evaluation. The underlying measurement system setup, retrieval strategy and sampling procedure are explained in the following.

System setup and retrieval strategy

The standard measurement system setup (STP) is simulated based on an intended upcoming ADL system for BL research. The simulator settings are set accordingly, an overview of the settings is given in Fig. C.1 with abbreviations provided in Table B.1. The aircraft characteristics are based on an unpressurized, medium-range turboprop aircraft, flying at the BL top at measurement speed under visual flight rules. Therefore, the aircraft is flying at an IAS of $65\,\mathrm{m\,s^{-1}}$ 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 scan pattern used in the ADLS study is founded in the scan geometry of existing systems (Weissmann et al., 2005b; De Wekker et al., 2012; Bucci et al., 2018). It consists of a 20-second full circle scan time (scan speed $18\,^\circ\,\mathrm{s^{-1}}$) at 60° elevation (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 10 Hz, with pulse width and range gate settings according to Sec. 4.2.

The standard wind profile retrieval strategy (RET) consists of u,v,w component retrieval using a volume-based profile separation. Retrievals are obtained from along-track averaging over X_a = 1300 m and the same across-track averaging distance of 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 1100 m above ground, it is adjusted accordingly if other scan elevations are used. The wind profile resolution in the vertical is chosen as 60 m, yielding one range gate within every vertical layer. Only wind profiles within the BL and entrainment zone are compared, as the ADLS study focuses on retrieval error under inhomogeneous, turbulent flow conditions. Therefore, the maximum wind profile altitude is set to 800 m, preventing the impact of larger scale features such as gravity waves in the free troposphere with longer de-correlation scales. For the standard retrieval the availability of all measurements is assumed, e.g. no data-loss due to blocking by clouds or other obstacles in the retrieval volume (so-called sector blanking in the following).

For analysis of system setup and retrieval strategy optimization a number of individual parameters adjustable in real-world measurements are varied according to the options detailed in Table B.1, while keeping the rest of the parameters at their standard setting.

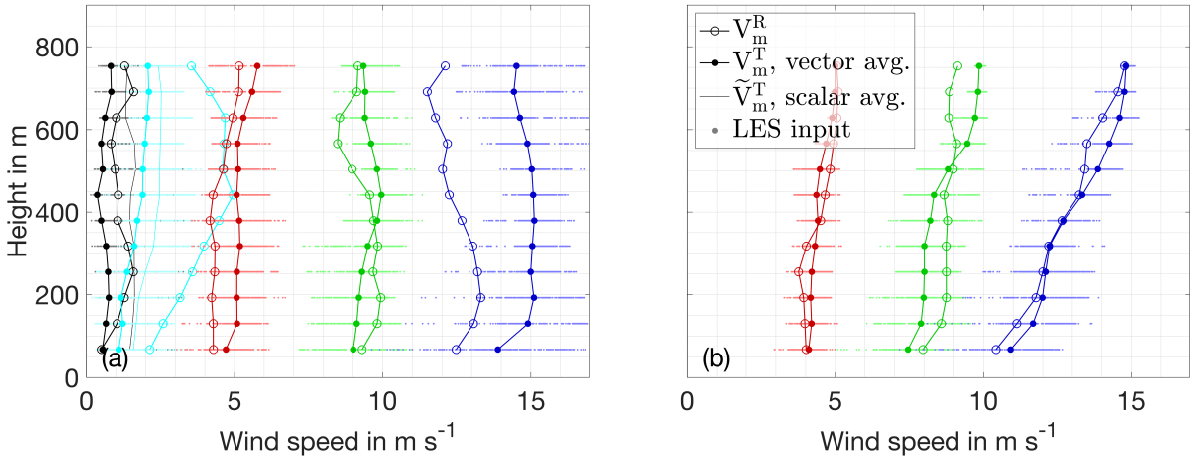

Fig. 4.3.: Examples of ADLS retrieved wind profiles and associated LES model truth for (a) LES set A and (b) LES set B using an ideal measurement system. The LES input denotes the LES wind speeds at all measurement locations visited by the lidar beam during the scan used to retrieve the individual profiles. The average of the LES input yields the LES model truth profile. For the $0\,\mathrm{m\,s^{-1}}$ and $2\,\mathrm{m\,s^{-1}}$ background wind cases the effect of scalar vs. vector averaging of the LES model input is shown. Retrieval errors are due to the violations of the flow homogeneity assumption used in AVAD. Examples of radial velocities and LSQ-fits used for AVAD retrieval are shown in Fig. 4.4.

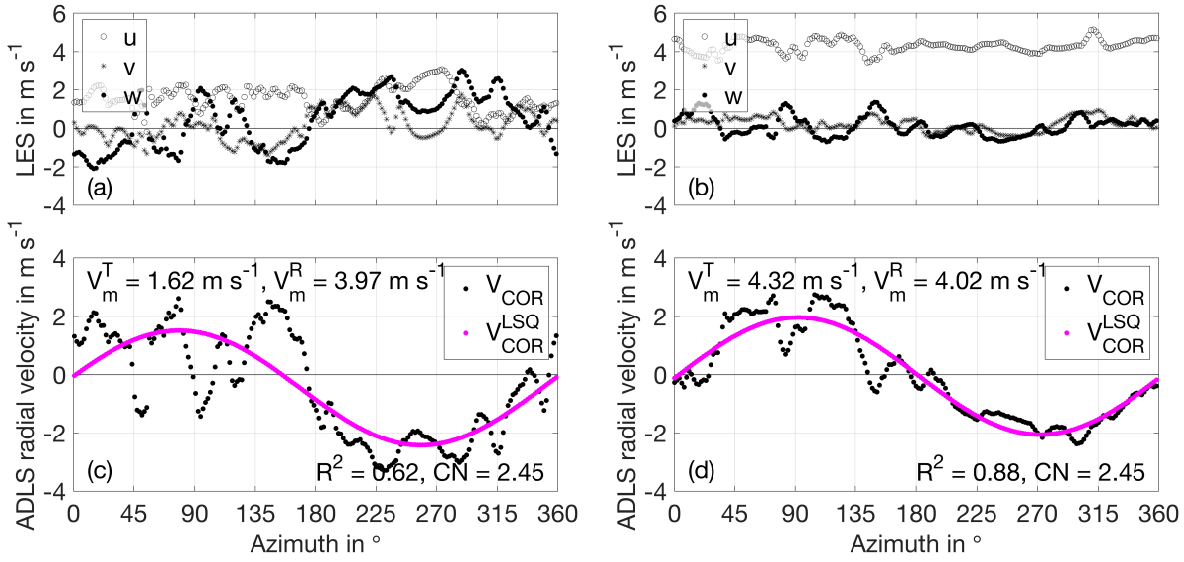

Fig. 4.4.: Illustration of the retrieval procedure using the AVAD technique for two of the wind profile points shown in Fig. 4.3 (at 320 m altitude). The LES input u, v, w values used to produce the simulated radial velocities are shown in (a) for the $2\,\mathrm{m\,s^{-1}}$ background wind case of LES set A and in (b) for the $5\,\mathrm{m\,s^{-1}}$ background wind case of LES set B. Panels (c) and (d) show the corresponding simulated motion corrected radial velocities, as well as LSQ-fits obtained from SVD-inversion used for AVAD retrieval. Please note that a movement towards the lidar is negative by convention, therefore 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.

Examples of retrieved profiles

Figure 4.3 shows examples of wind profiles which can be retrieved from the ADLS using the standard system setup and retrieval strategy. Therein V_m^T is the true wind speed, determined as the vector average of the input LES wind velocities (in the range gate volumes used to create the measurements, e.g. at the LES locations visited by the trace of the lidar beam along the scan circle). V_m^R is the ADLS retrieved wind speed from simulated AVAD wind profiling. Because an ideal measurement system is simulated, the errors identifiable in the retrieved wind profiles stem purely from the violation of the flow homogeneity assumption used for retrieval. Retrieval errors are clearly non-negligible and a number of further observations can be made from the profiles.

On average, as expected, retrieval errors are larger in LES set A compared to LES set B due to the stronger turbulence. In LES set A the observed retrieval error is so large that it can severely limit the usefulness of the retrieved wind profiles, especially at low wind speeds. In the case shown here the profile retrieved from the $2\,\mathrm{m\,s^{-1}}$ background wind case is indistinguishable from the one retrieved from the $5\,\mathrm{m\,s^{-1}}$ background wind case, despite the fact that the true input LES wind profiles are clearly different.

Further, the profiling error between neighboring vertical profile layers is not random but strongly correlated instead, because turbulent structures are correlated between vertically adjacent layers. This makes an individual wind profile appear smooth, despite non-negligible overall random error being present. The magnitude of the random error only becomes noticeable when looking at profiles sampled at different locations or times (Chap. 5). This disguise is likely one of the reasons why the wind profiling error due flow inhomogeneity has received little attention in literature so far.

Profiles where the retrieved wind speed is outside the range of input wind speeds can be observed in both LES sets. These are caused by strong supra-scan-volume contributions of the vertical wind, which is illustrated in Fig. 4.4 by showing the underlying LES input and simulated radial velocities. This analysis reveals the reason for the strongly offset retrieval from the $2\,\mathrm{m\,s^{-1}}$ background wind case of LES set A (Fig. 4.3a): The retrieval error is caused by an in-phase supra-scan-volume contribution of the vertical wind and resulting erroneous mapping. At the chosen scan elevation angle of 60 degree the projection of a $2\,\mathrm{m\,s^{-1}}$ vertical wind produces a radial velocity comparable to the projection of a $4\,\mathrm{m\,s^{-1}}$ horizontal wind. As a result, the LSQ-fit in Fig. 4.4c shows a very similar shape compared to the one in Fig. 4.4d, despite the different background wind speeds at hand. Hence, the retrieved wind speeds are also similar, resulting in a large retrieval error for the $2\,\mathrm{m\,s^{-1}}$ background wind case because the vertical

wind is erroneously mapped into horizontal wind. Additional superimposed sub-scan-volume flow inhomogeneity is also present in both measurements and results in noticeable deviations of the radial velocity from the LSQ-fit. A similar analysis holds for the $15\,\mathrm{m\,s^{-1}}$ background wind case of LES set A, but with an inverted sign of the supra-scan-volume vertical wind contribution, thereby leading to a negative offset of the retrieved profile (not shown).

Sampling procedure for statistical analysis

For statistical assessment of the errors associated with wind profiling in inhomogeneous flow conditions, instead of the few examples shown above, a large number of wind profiles are sampled from the LES domains in temporal and spatial proximity. Independence of the retrieved wind profiles is important to ensure a sufficiently large and independent data-set for robust analysis with reliable statistical properties. Therefore, correlation is minimized through an adequate sampling procedure explained in the following. Nevertheless, it should be noted that any potential correlation in wind profiling error does not affect the reported error levels systematically. Correlation reduces the number of effectively available wind profiles, and thereby influences the reliability with which the error levels are estimated, but it does not impact the error levels themselves. 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 technique). Due to the different characteristics of the LES data sets different sampling procedures are used which are explained in the following.
Note that despite investigating an overall independent error (between spatially and temporally separated profiles), the error of vertically adjacent wind profile points is not random for an individual profile due to the vertically correlated turbulence (see Fig. 4.3). This vertical correlation effect is present in both the ADLS and real measurements with equal magnitude.

Sampling procedure for LES set A

Five background wind fields are sampled using the time-frozen wind field sampling approach. Eight flight trajectories traverse the LES domain in parallel from south to north at a horizontal distance of 2600 m to each other, in a crosswind flight direction. Each of the eight trajectories is repeated seven times at 20-minute temporal spacing, giving a total number of 56 actual transects and ensuring temporal independence. To avoid correlation among the analyzed profiles from neighboring transects, all profiles are sampled from non-overlapping measurement domains in

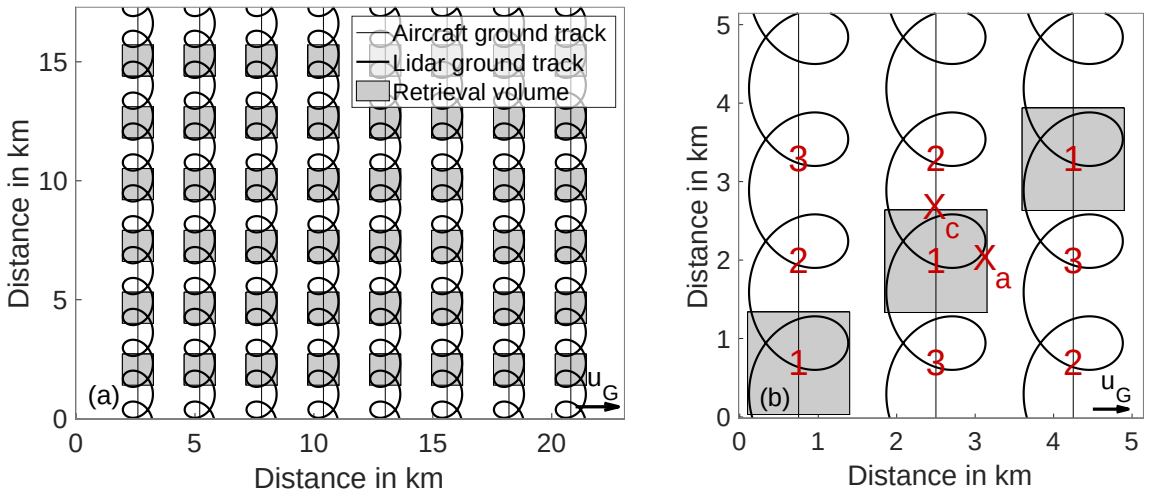

Fig. 4.5.: Illustration of the checkerboard technique used for wind profile retrieval in (a) LES set A and (b) LES set B. Shown are the aircraft transect locations, the lidar ground track as well as the extent of the retrieval volumes for the first time step (sampled simultaneously). For LES set A the flight trajectories and retrieval volumes are repeated every 20 minutes. For LES set B the flight trajectories are repeated every 1 minute and the retrieval volumes are shifted to the enumerated positions.

space, illustrated in Fig. 4.5a. To this end, the six retrieval volumes for each transect are also separated by an along-track distance of 2600 m.

Thereby, the resulting 48 retrieval volumes for 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 (O(100-1000 m)), ensuring statistical independence. A random offset within one retrieval volume length is added to all retrieval volumes within the same transect (not shown). The random start point prevents a systematic influence of the transect begin location on retrieval volume exploration for the different transects.

Sampling procedure for LES set B

Three background wind fields are sampled using the time-varying wind field sampling approach (the temporal wind field resolution used is 1 s). 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 Fig. 4.5b. 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, providing statistical independence.

The checkerboard pattern is shifted by X_a along the 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 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 (O(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.

An analysis of the spatio-temporal correlation of the obtained wind profile retrieval errors ΔV_m shows that the obtained errors are independent to a high degree for all background wind cases, time steps and retrieval altitudes (Fig. C.3, C.4, C.5). 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 technique. On the one hand, the autocorrelation values at lags $1, 2; 4, 5; 7, 8; \ldots$ reveal the spatial correlation of wind profiling error between neighboring retrieval volumes. On the other hand, the autocorrelation values at lag $3, 6, 9, \ldots$ 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. An example of weak structures in the alternating correlation coefficients is the slightly negative correlation at lag 1 for the $15\,\mathrm{m\,s^{-1}}$ background wind case at low altitudes (Fig. C.5b, d). 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.

In summary, both LES sets allow for a statistical analysis of AVAD wind profiling error in inhomogeneous flow conditions, which is conducted in the next chapter.

5. Assessing retrieval errors in airborne Doppler lidar wind profiling using virtual observations

In this chapter, the virtual ADLS observations, conducted according to the sampling strategy outlined in the previous chapter, are analyzed. The ADLS has the advantage that the input truth is known in full, which is impossible in real-world ADL measurements. Thereby, as an ideal measurement system without error in the system components is assumed, the previously inaccessible retrieval error in AVAD due to inhomogeneous flow conditions (in this case turbulence inside the BL) can be revealed. To this end, the virtual observations are used for a statistical analysis of errors in wind profile retrieval. Based on this, a measurement system setup and retrieval strategy optimization is presented for wind profiling in inhomogeneous flow. Further, the ADLS is used as a test-bed for validation of a new data-driven uncertainty estimation method. It is shown that the method allows for estimation of ADL wind profiling uncertainty for a wide range of atmospheric conditions, system setups and retrieval strategies.

5.1. Errors in wind profiling due to inhomogeneous flow conditions [1]

In the following, the wind profiling retrieval error $\Delta V_m = V_m^R - V_m^T$ is quantified statistically. As before, V_m^T is the true wind speed, determined as the vector average of the input LES wind velocities at the LES locations visited by the trace of the lidar beam along the scan circle. V_m^R is the ADLS retrieved wind speed from AVAD wind profiling. For statistical analysis, the mean absolute error (MAE) is used, which is given as:

$$\mathrm{MAE} = \frac{\Sigma_i^N |V_{m\,i}^T - V_{m\,i}^R|}{N}. \tag{5.1}$$

Here, N is the number of wind profile points fulfilling the quality control criteria. In later analysis, N_n is the normalized number of wind profile retrieval points, which is normalized with the number of retrieved wind profile points for the standard system setup and retrieval strategy.

[1] The results presented in this section are adapted from the publication: Gasch, P., Wieser, A., Lundquist, J. K. and Kalthoff, N., 2020: An LES-based airborne Doppler lidar simulator and its application to wind profiling in inhomogeneous flow conditions. Atmospheric Measurement Techniques, 13(3), 1609-1631.

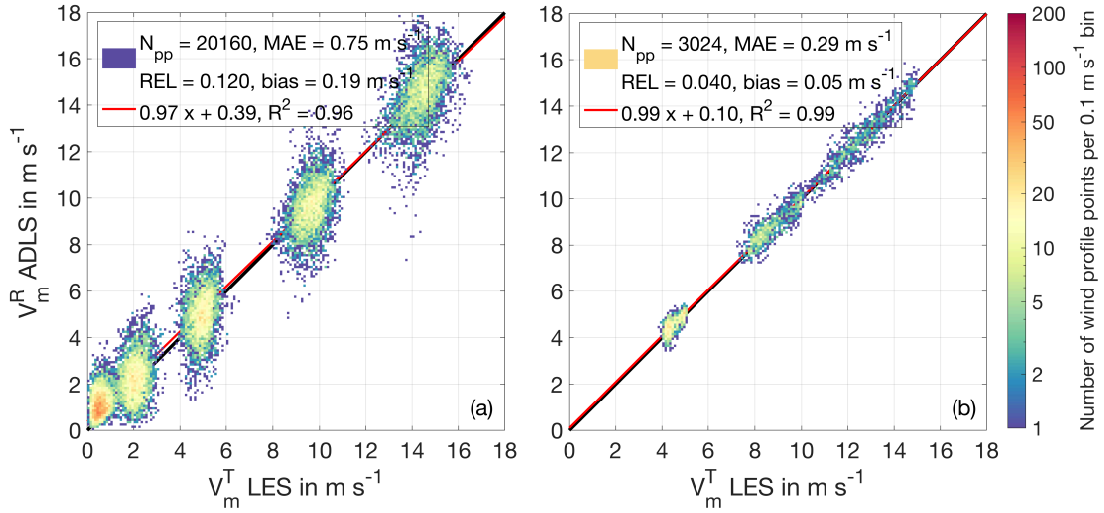

Fig. 5.1.: Comparison of LES truth and AVAD retrieved wind speed for an ideal measurement system for (a) LES set A and (b) LES set B. Color-coded is the histogram distribution of all measurements. The observable deviations from the 1:1 line reveal the AVAD retrieval error due to the violation of the flow homogeneity assumption. The system setup and retrieval strategy are the same between LES set A and B, more wind profiles can be retrieved from LES set A due to the larger simulation domain (Sec. 4.4.2).

Further, the relative root-mean-squared error (REL) is used in accordance with Guimond et al. (2014):

$$\mathrm{REL} = \sqrt{\left[\frac{\Sigma_i^N (V^T_{m\,i} - V^R_{m\,i})^2}{\Sigma_i^N (V^T_{m\,i})^2}\right]}. \tag{5.2}$$

The REL can provide additional information to the MAE 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 MAE and REL are driven by two factors, the random error and the bias of the retrieval. Therefore the bias of the retrieval is also reported as its average mean deviation,

$$\mathrm{bias} = \frac{1}{N}\sum_{i=1}^{N}(V^R_{m\,i} - V^T_{m\,i}). \tag{5.3}$$

5.1.1. Error quantification for standard system setup and retrieval strategy

Section 4.4.2 showed that a statistical analysis of error in wind profiling is possible using the given system setup, retrieval strategy and sampling procedure for both LES sets A and B. The wind profiling quality of the standard system setup and retrieval settings in inhomogeneous flow conditions can be evaluated from Fig. 5.1.

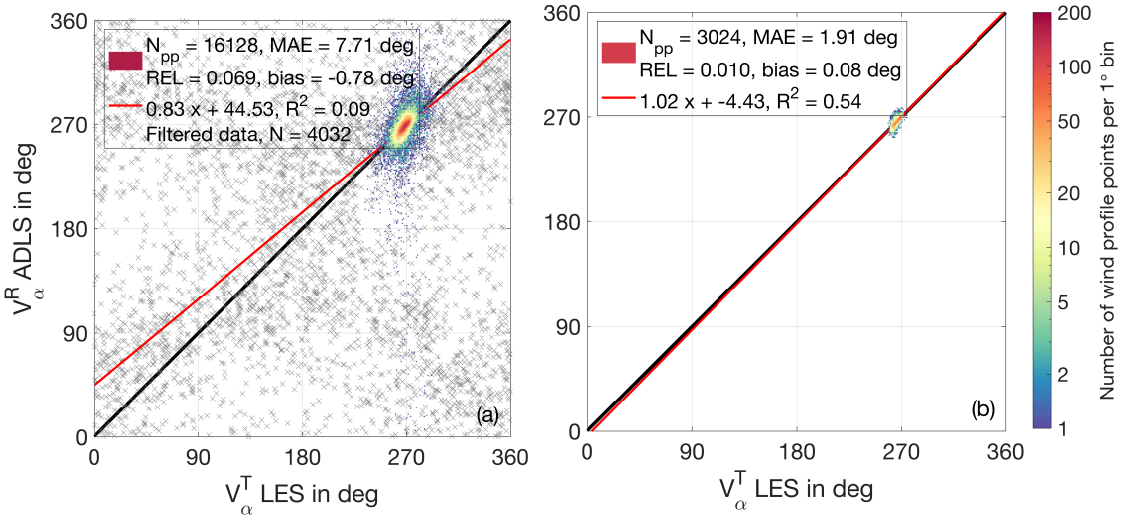

Fig. 5.2.: Same as Fig. 5.1 but for wind direction. The wind direction retrieval of the LES set A 0 m s^{-1} background wind case is excluded because the wind direction is not meaningful in this case.

The ADLS simulation using LES set A allows for retrieval of 1680 individual wind profiles, giving 20160 wind profile points for all altitudes (12 wind profile points per wind profile at the different retrieval altitudes). These consist of 336 wind profiles (4032 wind profile points) for each background wind case, yielding a sufficient statistical basis for evaluation. The 1680 wind profiles (20160 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 (see Sec. 2.3.1 for an overview). For example, 33 wind profiles (740 wind profile points) are compared to dropsonde data in Weissmann et al. (2005b), 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 ADLS simulation using LES set B allows for retrieval of 252 individual wind profiles giving 3024 wind profile points for all altitudes (12 wind profile points per wind profile at the different retrieval altitudes). These consist of 84 wind profiles (1008 wind profile points) for each background wind case, also yielding a sufficient statistical basis for evaluation.

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 flow inhomogeneity. Because the standard system setup and retrieval strategy are used, quality control with the *CN* does not remove any of the retrieved wind profile points. The results provide a number of interesting observations. Overall retrieval quality is high for both LES sets A, B with $R^2 = 0.96, 0.99$ and MAE $= 0.75, 0.29\,\mathrm{m\,s^{-1}}$. Yet, the violation of the AVAD flow homogene-

ity assumption is clearly non-negligible and can cause severe wind profiling error. Deviations above $3\,\mathrm{m\,s^{-1}}$ occur for LES set A, generated by strong BL turbulence. LES set B shows smaller retrieval error due to the reduced turbulence intensity, nevertheless, deviations above $1\,\mathrm{m\,s^{-1}}$ occur.

The wind speed retrieval is unbiased for the standard system setup and retrieval strategy investigated here, except for very low wind speeds and provided that R^2-based quality control is *not* applied. The retrieval bias at very low wind speeds is investigated further in Sec. 5.1.2 and so is the effect of R^2- and CN-based quality control. The effects of system setup and retrieval strategy on retrieval bias are investigated further in Sec. 5.2 and Sec. 5.3.

The LES cases with higher wind speeds show slightly increased mean absolute wind speed retrieval error, especially for the low turbulence LES set B. With decreasing wind speed (separately for each of the cases), associated with measurements in the BL under turbulent conditions, deviations remain unbiased (again, provided that R^2-based quality control is not applied).

The vertical distribution of wind profile error (Fig. C.6) mirrors that of the responsible BL turbulence (Fig. C.2). Relative errors are largest in the middle of the BL, where up- and downdrafts have maximum intensity. Towards the ground, a reduction in wind profiling error is observable for all background wind cases. There, the size of the turbulent elements becomes smaller and consequently the lidar scan averages over more eddies, thereby better fulfilling the flow homogeneity assumption. Towards the top of the BL, 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 background wind cases. In the homogeneous flow of the free atmosphere the wind profile error vanishes (not shown).

The wind direction retrieval is of similar quality compared to the wind speed retrieval (Fig. 5.2). For the $0\,\mathrm{m\,s^{-1}}$ background wind case of LES set A the wind direction is not a meaningful measure, it is thereby excluded before analysis. As expected, higher scatter is observed for the LES set A compared to LES set B due to the more intense turbulence. Retrieved wind profile directions cluster in an area around 270° for both LES sets, representing the westerly wind direction. The wind direction retrievals for the lower background wind case scatter more than for higher background wind speeds, as the wind direction becomes increasingly difficult to retrieve. The retrievals exhibits slightly degraded quality criteria (lower R^2, small y-axis intercept) 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.

In summary, based on the results discussed so far, it is concluded that ADL wind profiling error due to violation of the flow homogeneity assumption can be of non-negligible magnitude. The errors reported in this work, 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., 2005b; De Wekker et al., 2012; Bucci et al., 2018). Wind profiling error due to flow inhomogeneity must therefore not be neglected as an important source of error when assessing the measurement accuracy of ADL systems, especially for measurements inside the BL.
In the next two sections, the performance of the most commonly used quality control criteria R^2 and CN is evaluated (Sec. 3.3), specifically their relation to wind profile error and their adequacy in detecting violations of the flow homogeneity assumption and collinearity in the model geometry.

5.1.2. Applicability and reliability of R^2-based quality control

In this section, a performance check of the R^2-based quality control is conducted for the standard system setup and retrieval setting. The analysis is conducted only for LES set A, but the effects discussed are also present in LES set B with reduced magnitude.
ADLS results for the ideal measurement system setup and standard retrieval setting present a number of interesting findings (Fig. 5.3). The majority of the R^2 values is located above levels of 0.9, but with decreasing wind speeds lower values of R^2 also occur due to the increasing influence of turbulence relative to the mean wind. Even at high $R^2 > 0.90$, significant wind profile errors above $3\,\mathrm{m\,s^{-1}}$ do exist. The CN values cluster in a narrow range with values between $3-4$. CN-based quality control is not applicable here as the sampling volume is generally well-explored by the lidar for all retrievals.
A problematic, previously undocumented R^2 feature occurs at low wind speeds, when the turbulence intensity becomes large compared to the background wind speed. In this region, a clear dependence of retrieval error on the R^2 value is observable (Fig. 5.3a, b, c). With increasing R^2, the bias of the retrieved wind speed increases linearly to values in the range $1-2\,\mathrm{m\,s^{-1}}$. This problem exists despite the fact that the individual retrieved components themselves are unbiased (Fig. C.7). The observed correlation of wind profile retrieval error and R^2 prohibits any quality control of the wind speed retrieval using R^2, as a bias can be introduced thereby.
The observed behavior is caused by the lidar measurement method: at low wind speeds, the radial velocity and thereby retrieval is strongly influenced by the vertical wind and less by the horizontal wind, especially at steep elevation angles (see also Fig. 4.4). 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 map more into horizontal wind, causing a stronger positive bias. The values of the horizontal wind speed in the range of $0-3\,\mathrm{m\,s^{-1}}$ correspond to a mean vertical wind of $0-1.5\,\mathrm{m\,s^{-1}}$ amplitude being mapped into

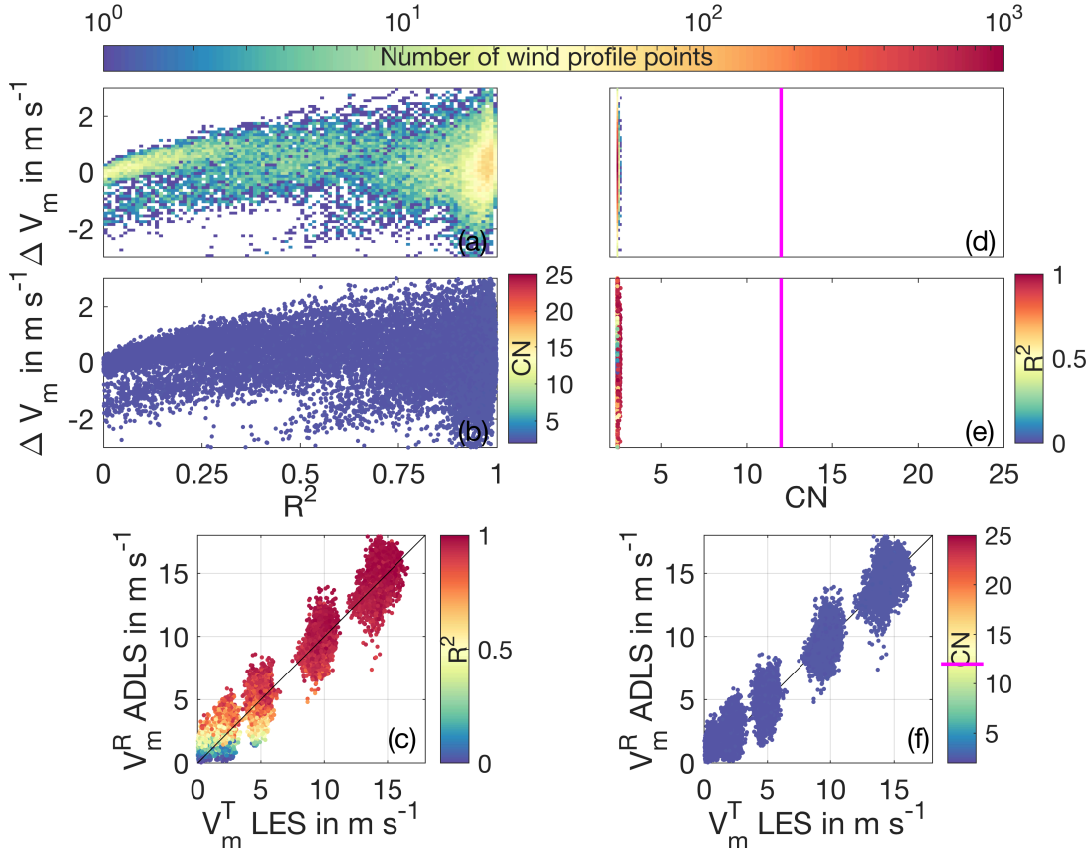

Fig. 5.3.: Quality control criteria for the standard system setup and retrieval strategy for all background wind cases of LES set A. a) Color-coded histogram of number of occurrence for retrieval error and R^2. b) Retrieval error and R^2, color-coded is CN. c) LES truth and ADLS retrieved wind speed, color coded is R^2. d) Same as a), but for CN. e) Retrieval error and CN, color coded is R^2. f) Same as c), but for CN. The magenta line indicates the CN quality control threshold applied.

the horizontal wind at 60° elevation. At different background wind speeds the R^2 correlation curves exhibit different slopes (the distinct clusters in Fig. 5.3a, b are associated with the different background wind cases). For the $0\,\mathrm{m\,s^{-1}}$ background wind case the slope is not centered around $0\,\mathrm{m\,s^{-1}}$ but elevated to positive values instead, because the overall retrieval is biased in this case (the effect is similarly present for the $2\,\mathrm{m\,s^{-1}}$ background wind case but at reduced magnitude).

The noisy (sub-scan-volume and supra-scan-volume) w variations can overwhelm the smaller horizontal wind signal. Thereby, the coefficient of determination is correlated with the degree of wind profiling error introduced. Due to the systematic behavior, a bias can be introduced in the retrieval if R^2-based quality control is applied. Unfortunately, the systematic dependence cannot be easily corrected, as the a priori wind speed is unknown in real-world measurements. Thereby, the exact slope and location of the R^2 correlation curve is unknown (Fig. 5.3a). In line

with this, a closer look at the higher background wind cases reveals a similar dependence of wind speed retrieval error on the R^2 level at higher wind speeds as well (Fig. C.8c). Or worded differently: The range and slope of the R^2 spread is a function of the underlying ratio between background wind speed and turbulence intensity. As the background wind speed is unknown prior to ADL retrievals, the R^2 correlation bias cannot be easily corrected and R^2-based quality control should be avoided.

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 $-1\,\mathrm{m\,s^{-1}}$ 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, it has to be assessed that the R^2, commonly assumed to capture the degree of turbulence and thereby flow homogeneity assumption violation, fails to do so reliably. Especially for low wind speeds the application of a R^2 quality control threshold mainly filters for smoother in-phase variations of the vertical wind. Consequently, R^2-based quality control can cause a retrieval bias, an important finding that is not expected and, to my knowledge, undocumented so far. Especially at low wind speeds, the R^2 is not an appropriate quality control criteria. A possible bias for unknown reasons, varying by day, has been reported by Weissmann et al. (2005b) 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 retrieved 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), Pu et al. (2010) (airborne; Fig. 1b), Chouza et al. (2016a) (airborne; Fig. 2, Fig. 3b under turbulent conditions inside the BL) and Pauscher et al. (2016) (ground-based; Fig. 5e).

Consequently, this work presents and justifies a new restriction on wind speed retrieval using Doppler systems: in order for R^2-based quality control and wind speed profiles to be unbiased, the radial contribution of the vertical wind deviation magnitude has to be small compared to the horizontal wind magnitude. If this criterion is violated, the mapping of vertical wind into horizontal wind can bias the wind speed retrieval. Therefore, the following procedure is recommended for wind profile analysis: In doubtful situations, where the approximate magnitude of the horizontal and/or vertical wind 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 horizontal wind can be analyzed reliably, although at coarse spatial resolution. If the outcome of this analysis yields a horizontal wind magnitude much above the vertical wind magnitude (e.g. by an order of magnitude), a further analysis at higher spatial resolution can be conducted. If, on the other hand, the magnitude ratio is below this threshold (e.g. the vertical wind is non-negligible), the spatial resolution of the retrieval cannot be refined further without extra precautions, due to the possibility of erroneous mapping. In doubtful situations it is better to conduct the wind profiling analysis based on the individual wind components instead of the wind speed calculated from them. The retrieved wind components remain unbiased even at low wind speeds because the erroneous vertical wind mapping results in a vector with systematically too large magnitude but random orientation. If advanced and flexible measurement system are available, it is possible to adjust the measurement system setup and retrieval strategy based on the encountered flow conditions. To facilitate this optimization and minimize the error introduced by inhomogeneous flow conditions, an analysis of preferable measurement system setups and retrieval strategies is conducted in Sec. 5.2 and 5.3.

5.1.3. Applicability and reliability of CN-based quality control

In this section, a performance check for quality control using the condition number CN is conducted. In order to do so, the high resolution along-track retrieval setting (AVG = 650 m) is used along with the standard system setup and retrieval strategy otherwise. For the high resolution retrieval the analyzed volumes show varying coverage by the simulated ADL, which increases the CN spread and allows for assessment of CN-based quality control reliability. Figure C.9 gives an overview of the achieved retrieval quality for an along-track averaging distance of $X_a = 650\,\mathrm{m}$ and with $CN < 12$ quality control criteria applied. As expected, the maximum number of retrievable profile points is doubled, however, a much larger potential retrieval error exists. CN thresholding excludes the unreliable retrievals, thereby the retrieval error is at levels comparable to the previous analysis, but with more profile points available (LES set A: N_{pp} = 28622, MAE = 0.93 $\mathrm{m\,s^{-1}}$, bias = 0.31 $\mathrm{m\,s^{-1}}$; LES set B: N_{pp} = 3906, MAE = 0.45 $\mathrm{m\,s^{-1}}$, bias = 0.08 $\mathrm{m\,s^{-1}}$). The following detailed analysis is conducted only for LES set A but the effects are also present in LES set B, at reduced error magnitude.

Figure 5.4 shows the same analysis scheme as presented before. As discussed, much more retrieval scatter exists due to the challenging retrieval conditions and associated larger CN spread (Fig. 5.4d).

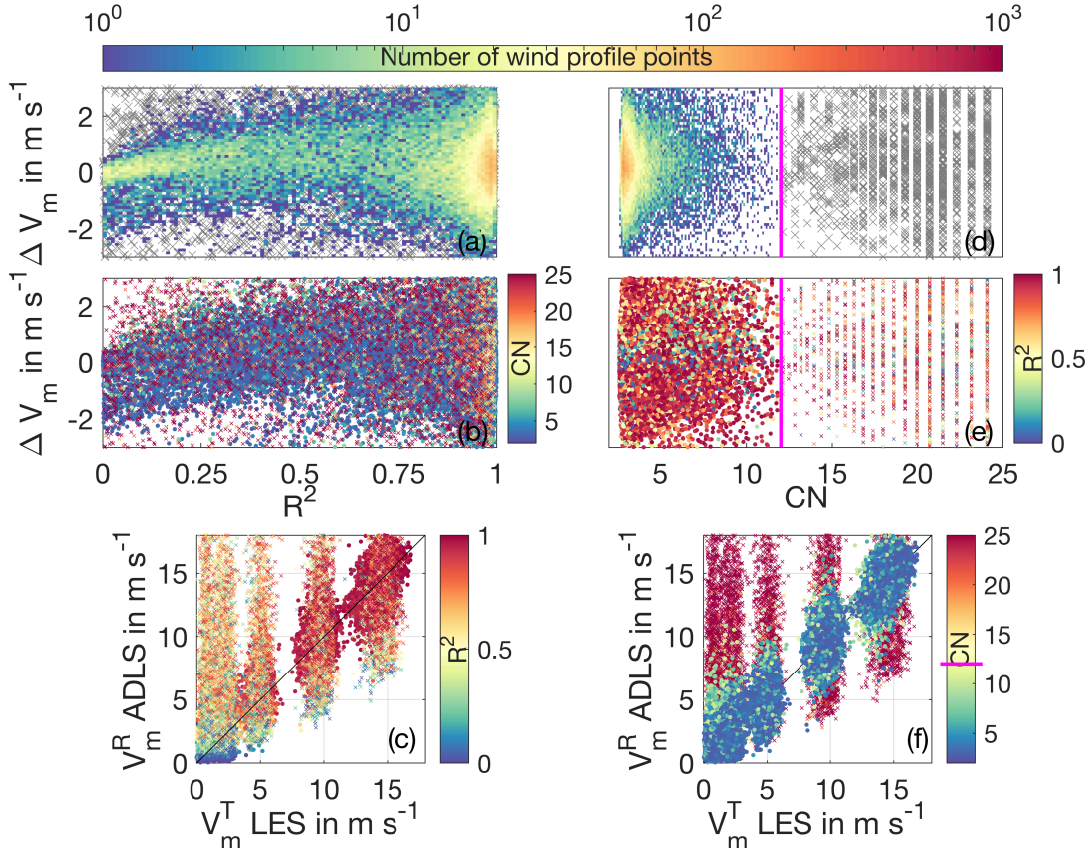

Fig. 5.4.: Quality control criteria for the standard system setup and challenging AVG = 650 m retrieval strategy for LES set A. a) Color-coded histogram of number of occurrence for retrieval error and R^2. Gray crosses show values eliminated by quality control. b) Retrieval error and R^2, color-coded is CN. c) LES truth and ADLS retrieved wind speed, color coded is R^2. Profile points which pass quality control 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. The magenta line indicates the CN quality control threshold applied.

The R^2 distribution shows a similar, problematic behavior as discussed in the previous section. Further, no decrease in retrieval error with higher R^2 values is observed (Fig. 5.4c). However, the results demonstrate that CN-based quality control works reliably, independent of the encountered flow situation. Retrievals with large error are related to high values of CN (Fig. 5.4f). Except for the expected increase, no systematic correlation or bias of the retrieval error with the CN is observed (Fig. 5.4d), as desired. On the one hand, further lowering the CN threshold improves retrieval quality slightly, albeit at the cost of further reducing the number of retrievable wind profile points. For example, with a $CN < 5$ threshold, the achieved quality metrics are for LES set A: N_{pp} = 23389, MAE = 0.82 m s^{-1}, bias = 0.23 m s^{-1}; and for LES set B: N_{pp} = 3325, MAE = 0.36 m s^{-1}, bias = 0.02 m s^{-1}. On the other hand, setting a higher CN threshold degrades retrieval quality but allows for retrieval of more wind profile points. For ex-

ample, with a $CN < 19$ threshold, the achieved quality metrics are for LES set A: $N_{pp} = 29608$, MAE = 0.98 m s^{-1}, bias = 0.35 m s^{-1}; and for LES set B: $N_{pp} = 3993$, MAE = 0.47 m s^{-1}, bias = 0.08 m s^{-1}. Beyond a CN threshold of 20 retrieval errors grow rapidly. A threshold value of $CN < 12$ is therefore a good compromise with some safety margin included, it is thereby recommended for real-world measurements.
In summary, quality control using the commonly applied R^2 threshold criteria is shown to be non-advisable by the ADLS results. R^2-based quality control does not detect violations of the flow homogeneity assumption reliable and can bias the wind speed retrieval. CN-based quality control on the other hand is reliable and recommended without limitations.

5.2. System optimization for wind profiling in inhomogeneous flow

It is desirable to obtain a better handle on the non-negligible wind profiling error due to inhomogeneous flow conditions in order to reduce its effect as much as possible. Therefore, this work utilizes a threefold approach in the following: First, it is investigated to what extent wind profiling error can be reduced through an appropriate system setup in this section. Second, the next section investigates the influence of the retrieval strategy on wind profiling error. Third, even for the most adequate system setups and retrieval strategies presented based on these results, wind profiling error cannot be fully mitigated. Therefore, the possibility of a data-driven uncertainty estimation based on the measured data is investigated in Sec. 5.4.
In the following first step, three system setup parameters which are potentially available for optimization are varied to investigate their influence on wind profiling quality. The parameters are the scan elevation angle and scan speed as well as the lidar measurement frequency.

Scan elevation angle

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. 5.5a, b). 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, but the flow homogeneity assumption is not necessarily fulfilled better. Consequently, at elevations steeper than

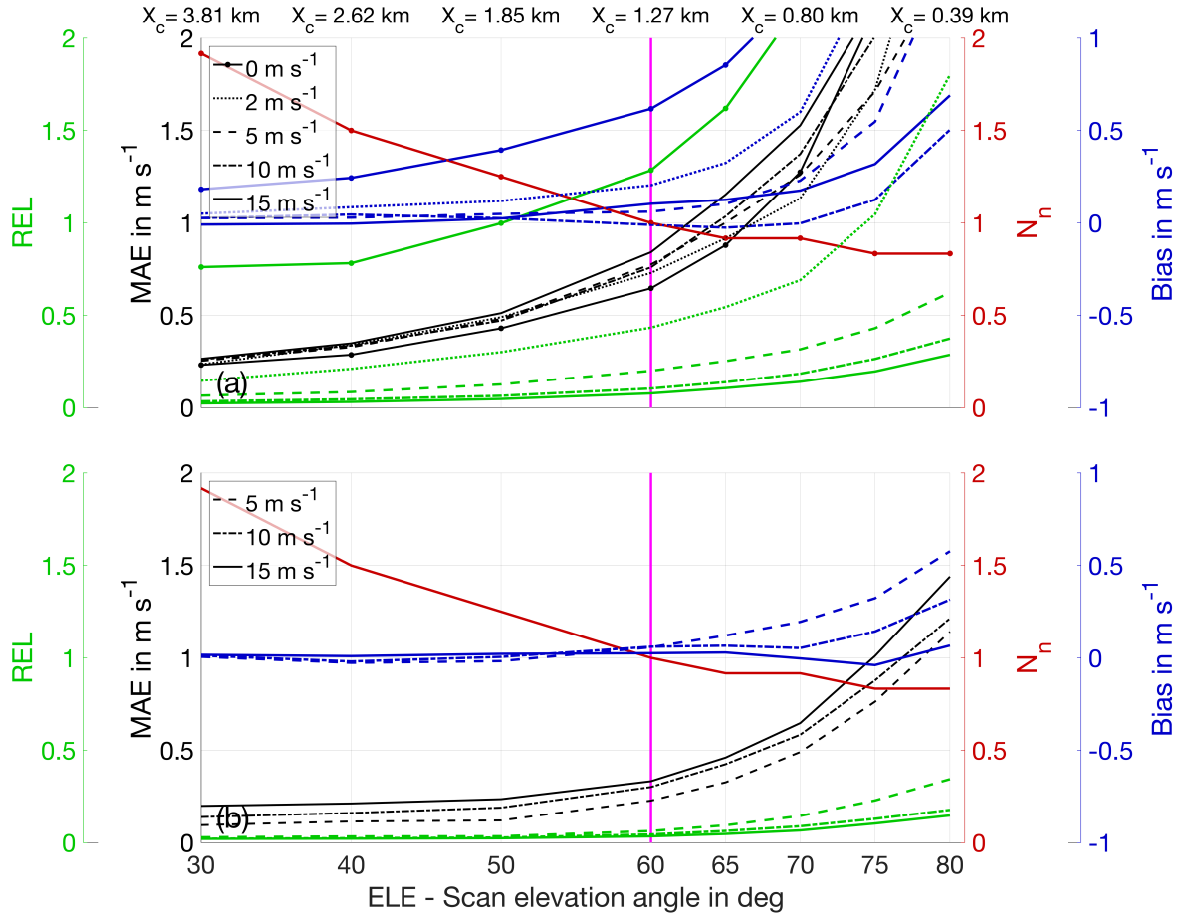

Fig. 5.5.: REL, MAE, N_n and bias for six scan elevation angles using STP and RET otherwise. a) LES set A with high surface sensible heat flux and five background wind cases. b) LES set B with low surface sensible heat flux and three background wind cases. Along-side the scan elevation angle the across-track averaging distance X_c is changed to the values specified on top in order to accommodate for the distance covered by the lidar beam. The scan elevation angle of the standard system setup is indicated as a magenta line.

60° the wind profiling quality is degraded, visible through the increased MAE (which cannot be improved through quality control). Qualitatively the LES sets A and B show a very similar MAE development with increasing scan elevation. However, the observed error levels differ by a factor of approximately two, highlighting the importance of the atmospheric conditions present, as discussed before (Sec. 5.1.1). The degradation in retrieval quality with increasing scan elevation is not just due to a larger retrieval scatter, in addition the retrieval also becomes biased due to erroneous mapping of vertical wind into horizontal wind. The bias is enabled by the large radial velocity contribution of the vertical wind and consequently increases with increasing vertical wind magnitude. Therefore, LES set A shows stronger retrieval bias than LES set B. The bias starts to develop for low wind speeds at more shallow elevation angles already, but eventually encompasses all analyzed background wind cases. The retrieval from the $0\,\mathrm{m\,s^{-1}}$

background wind case is biased for all elevation angles. Despite the bias, this case shows the lowest MAE for some scan elevation angles, which is due to the reduced retrieval scatter because negative wind speed values are not retrievable (see Fig. 5.1). R^2-based quality control worsens the pre-existing wind profiling bias and can introduce bias into previously unbiased retrievals for all background wind cases (not shown, see also Sec. 5.1.2). Hence, an R^2-based quality control is strongly recommended against once again. As expected, for steeper elevation angles the number of usable wind profile points decreases due to the increased projection of the range gate length onto the vertical. Thereby, the maximum number of available wind profile points is lower. In summary, for the standard retrieval strategy, scans steeper than 60° 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 less defined measurement footprint. On the one side, ADLS results show decreased MAE levels and bias, as well as an increased number of wind profile points being available. The decrease in MAE level and bias is due to a smaller influence of the vertical wind. Nevertheless, without background wind speed the retrieval can remain biased, which is an undocumented problem to my knowledge. The increased number of retrieved wind profile points is due to the more shallow scan elevation, which leads to a smaller projection of the range gate length onto the vertical. The additional points do not necessarily contain additional, independent information due to the vertical correlation of turbulence and thereby wind profile error, limiting the benefit of the increased 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.

In summary, the results obtained here allow for exemplary system optimization for wind profiling in inhomogeneous flow. Generally, atmospheric situations with higher turbulence or lower wind speeds are more problematic. If possible, the system setup should be flexible to allow accommodation for different atmospheric conditions. For example, the reduction in wind profiling error between LES set A and LES set B at 60° scan elevation is as large as the improvement that is achieved by scanning at 40° compared to 60° in LES set A. Scan elevations steeper than 60° should be avoided if inhomogeneity of the atmospheric flow is expected, especially at low wind speeds, because the resulting ADL retrieval can be biased.

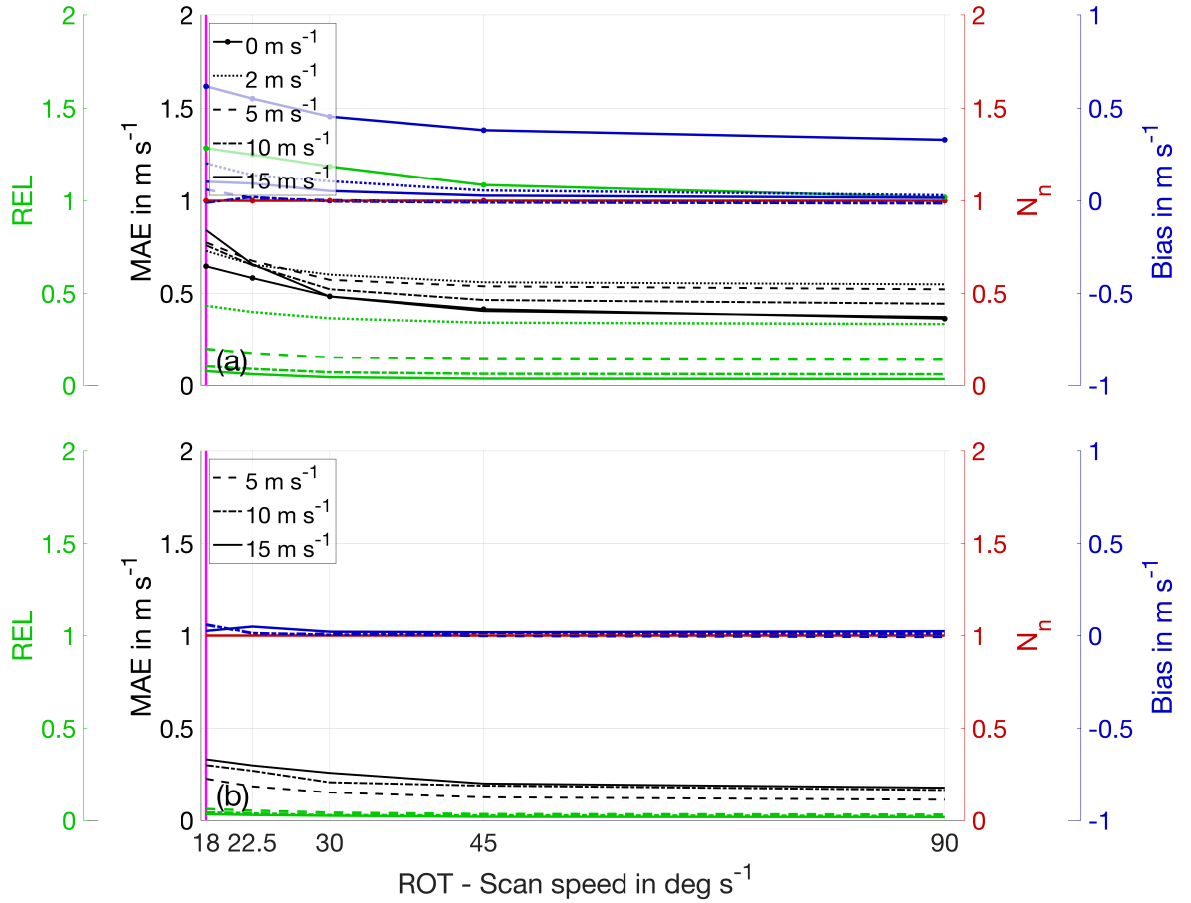

Fig. 5.6.: REL, MAE, N_n and bias for five scan scan speeds using STP and RET otherwise. a) LES set A with high surface sensible heat flux and five background wind cases. b) LES set B with low surface sensible heat flux and three background wind cases. The scan speed of the standard system setup is indicated as a magenta line.

Scan speed

Scan speed (ROT) also has a noticeable impact on wind profiling quality (Fig. 5.6), although it is less pronounced than that of the scan elevation angle. As before, the behavior is qualitatively similar between LES set A and B but at different error levels. At slower scan speeds the number of independent measurements in the retrieval volume is reduced due to correlation among the measurements (the number of measurements is the same as measurement frequency is unchanged). Faster scan speeds allow for greater azimuth diversity and reduced correlation among the measurements. Thereby, an improved wind profiling quality with lower MAE is possible, despite having an overall equal number of measurement points available in each volume. All scan speeds achieve the same number of retrievable wind profile points. For faster scan speeds the retrieval bias at low wind speeds is reduced, because the retrieval volume is better explored. Therefore, less potential for erroneous mapping of vertical wind exists, making faster

scan speeds preferable. However, the retrieval error eventually approaches a limit, as for very fast scan speeds the scan volume is completely explored. Thereby, the effect of the supra-scan-volume contribution cannot be reduced further because the horizontal shear of the vertical wind at retrieval volume scale is fully captured and unavoidable. Further, for very fast scans the lidar measurements from subsequent rotations become correlated in along-track direction. This is the case when the scan speed is so high that the distance traveled by the aircraft between subsequent scan rotation is less than the integral length scale of the flow. For the system setup investigated here, this is approximately the case for scan speeds $> 60\,^{\circ}\,\mathrm{s}^{-1}$ when the along-track distance between subsequent scans becomes $< 390\,\mathrm{m}$. It is also for this reason that higher background wind cases show a stronger reduction with increasing scan speed. For higher wind speeds, the integral length scales are reduced, thereby the degree of correlation from spatially close lidar beam paths is reduced. An important aspect which is not captured by the ADLS is the influence of increasing scan speeds 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.

Lidar measurement frequency

ADLS results show that an increased lidar measurement frequency (FME) does not improve wind profiling quality (Fig. 5.7). This insensitivity exists because higher measurement frequencies produce measurements in between the lower frequency measurements which are highly correlated. Due to the correlation (resolving the turbulent structures in greater detail but not fulfilling the flow homogeneity assumption better) the additional measurements do not provide new information that can be used to improve the retrieval quality. 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.
The different lidar measurement frequencies are an important test environment for the data driven uncertainty estimation method developed in Sec. 5.4. There, the degree of correlation present in the measurements needs to be estimated reliably to produce an accurate uncertainty estimate independent of the lidar measurement frequency.

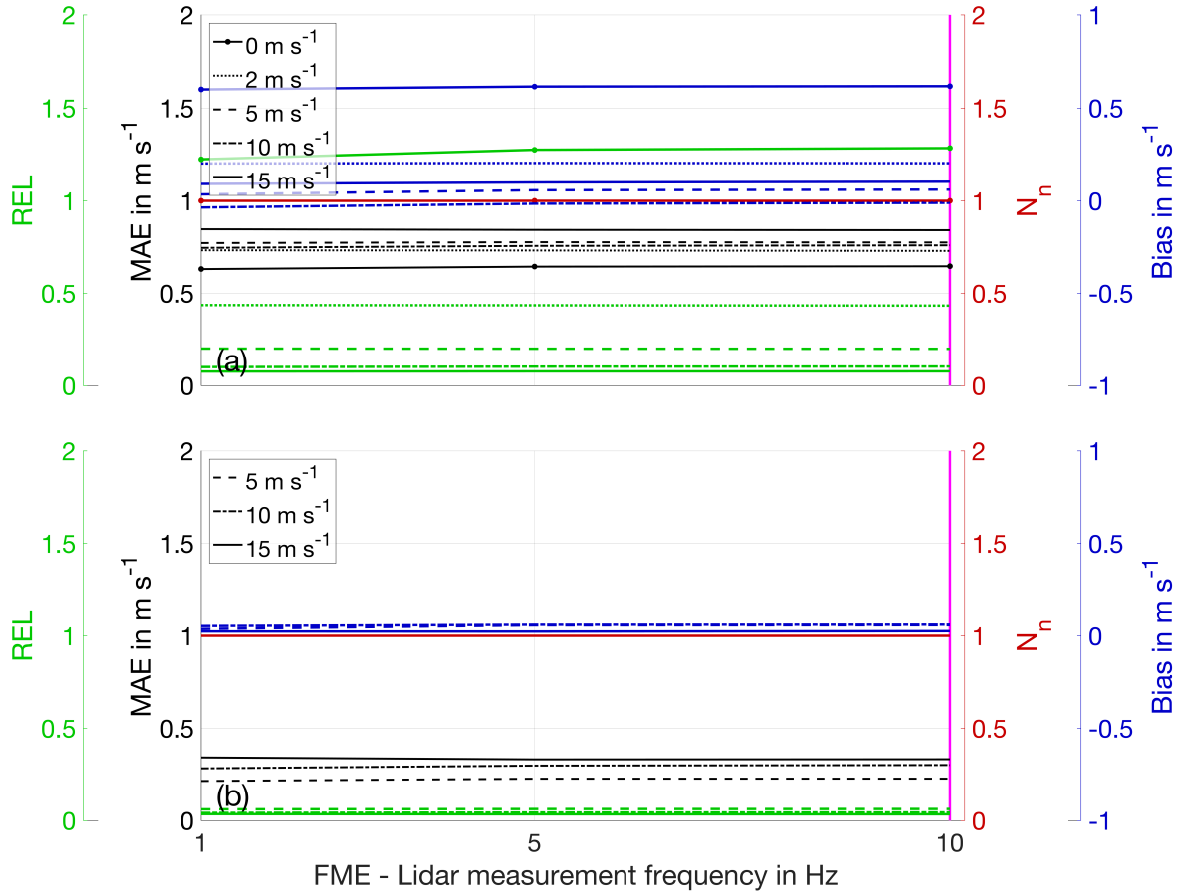

Fig. 5.7.: REL, MAE, N_n and bias for three lidar measurement frequencies using STP and RET otherwise. a) LES set A with high surface sensible heat flux and five background wind cases. b) LES set B with low surface sensible heat flux and three background wind cases. The lidar measurement frequency of the standard system setup is indicated as a magenta line.

5.3. Retrieval optimization for wind profiling in inhomogeneous flow

As a VVP type of algorithm (Waldteufel and Corbin, 1978; Boccippio, 1995) is applied to retrieve the wind speed, a number of options exist to vary the retrieval settings. In this section, the impact of three retrieval parameters, which are available for adjustment, are varied to investigate their influence on wind profiling quality. The parameters are the along-track averaging distance and the vertical extent of the volume used for the retrieval. Further, the impact of data availability on inversion stability is investigated. This investigation is done through omission of defined sectors in the retrieval volume in a procedure termed sector blanking. Thereby, the presence of obstacles in the retrieval volume, e.g. clouds or mountains, in real measurements is simulated.

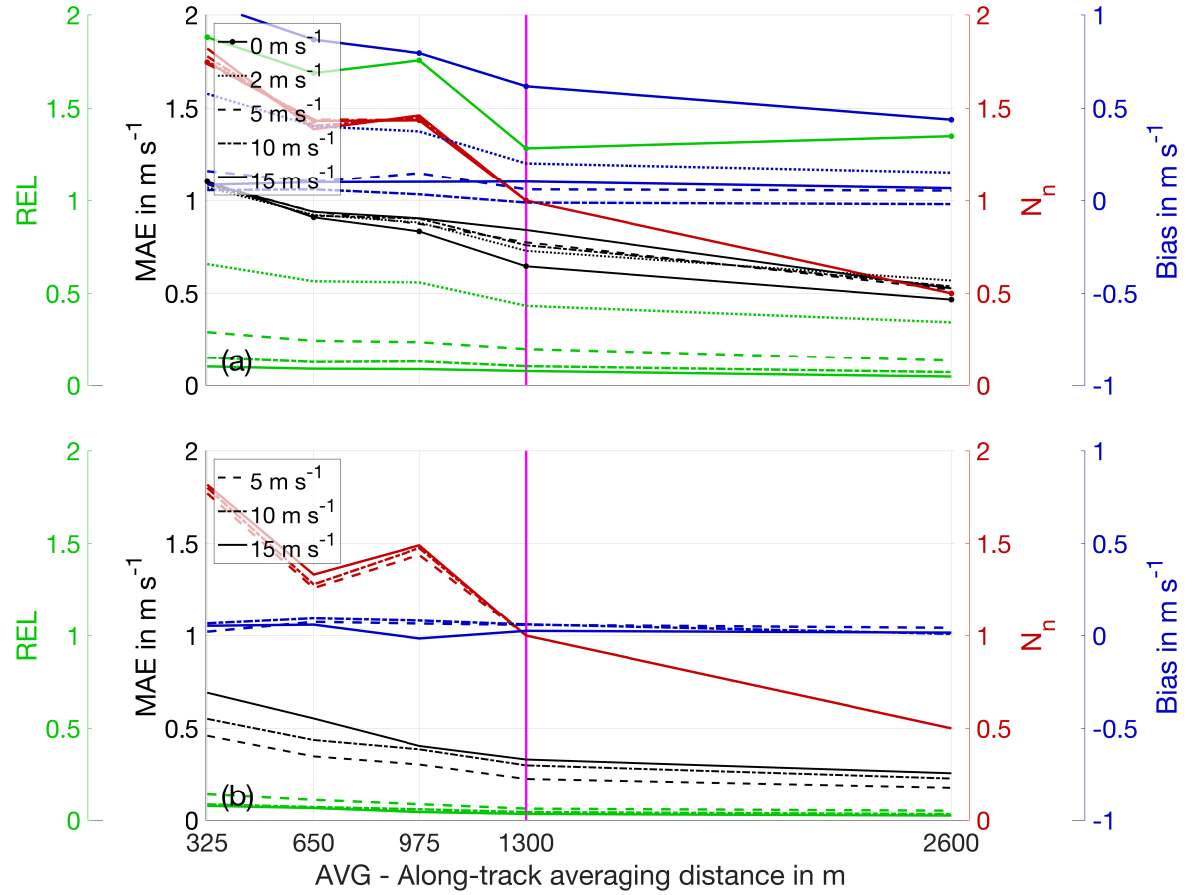

Fig. 5.8.: REL, MAE, N_n and bias for three along-track averaging distances using STP and RET otherwise. a) LES set A with high surface sensible heat flux and five background wind cases. b) LES set B with low surface sensible heat flux and three background wind cases. The along-track averaging distance of the standard retrieval strategy is indicated as a magenta line.

Along-track averaging distance

For the standard system setup and retrieval strategy the along-track averaging distance (AVG) is varied between 325-2600 m (determining the along-track wind profiling resolution). The along-track distances correspond to the approximate distance covered by the aircraft during 0.25-2 scanner rotations. However, because a volume based retrieval strategy is used, the moving aircraft allows for directing the laser beam forward and aft of the aircraft and therefore measures the same spot at different times under varying azimuth angles.

As expected, varying the along-track averaging distance used for profile retrieval has a noticeable influence on wind profile error (Fig. 5.8). High-resolution wind profiling at 325 m is associated with a larger MAE and a slightly more biased retrieval in the case of LES set A. Nevertheless, *CN*-based quality control manages to keep the error at acceptable levels, but also reduces the number of available profile points, as is discussed in Sec. 5.1.3. Less than double

the wind profiles can be retrieved despite the quadrupled along-track resolution. The decreased number of retrievable profile points compared to the theoretical maximum is driven by a decrease at higher altitudes (not shown). At greater altitude, the lidar beam distance from the aircraft is smaller and hence the before and aft extension of the lidar beam are smaller and the distance traveled by the lidar beam is shorter. Therefore, a number of 'mismatching' retrieval volumes exist where no sufficient coverage of the lidar beam is achieved because no separate observations from two separate aircraft positions are available. Hence, the number of retrievable profile points is higher close to the ground, where it is the most interesting as the most rapid changes in the wind field are expected. Retrieval of more wind profile points also at higher altitudes can be enabled through shallow elevation angles and/or faster scan speeds.

Doubling the along-track resolution compared to the standard retrieval strategy (AVG = 650 m) results in an increase of available wind profile points of less than 50%. The error levels are comparable to that of the AVG = 325 m case, slightly elevated from the standard retrieval strategy. The slight difference in the number of retrieved profile points between LES set A and B is caused by the slightly different sampling strategies applied (Sec. 4.4.2).

An interesting effect occurs for the AVG = 975 m case, where an increase in retrievable wind profile points compared to the AVG = 650 m case is observed for both LES set A and B. For this retrieval strategy, almost the maximum number of theoretically available profile points can be retrieved (1.5 times the points compared to the standard). CN-based quality control removes very few points as the retrieval volumes are adequately explored at all altitudes. The error levels are slightly elevated compared to the standard retrieval strategy, but below those of the AVG = 650 m case.

The error decreases for all background wind cases if the standard averaging distance of 1300 m is used. The decrease, however, is not linear and a further increase of the averaging distance to 2600 m does not yield an improvement of similar magnitude. It does, however, reduce the bias of the lowest background wind cases slightly because the potential for erroneous mapping due to supra-scan-volume flow inhomogeneity is reduced. As expected, this comes at the cost of having only half the number of profiles available.

Overall, the results demonstrate the potential for improved along-track resolution when using adequate system setups and retrieval strategies. Further, the above analysis can also be inverted, that is, if an acceptable error threshold is given the required along-track averaging distance can be determined for given atmospheric conditions, measurement system setup and retrieval strategy.

Vertical averaging distance

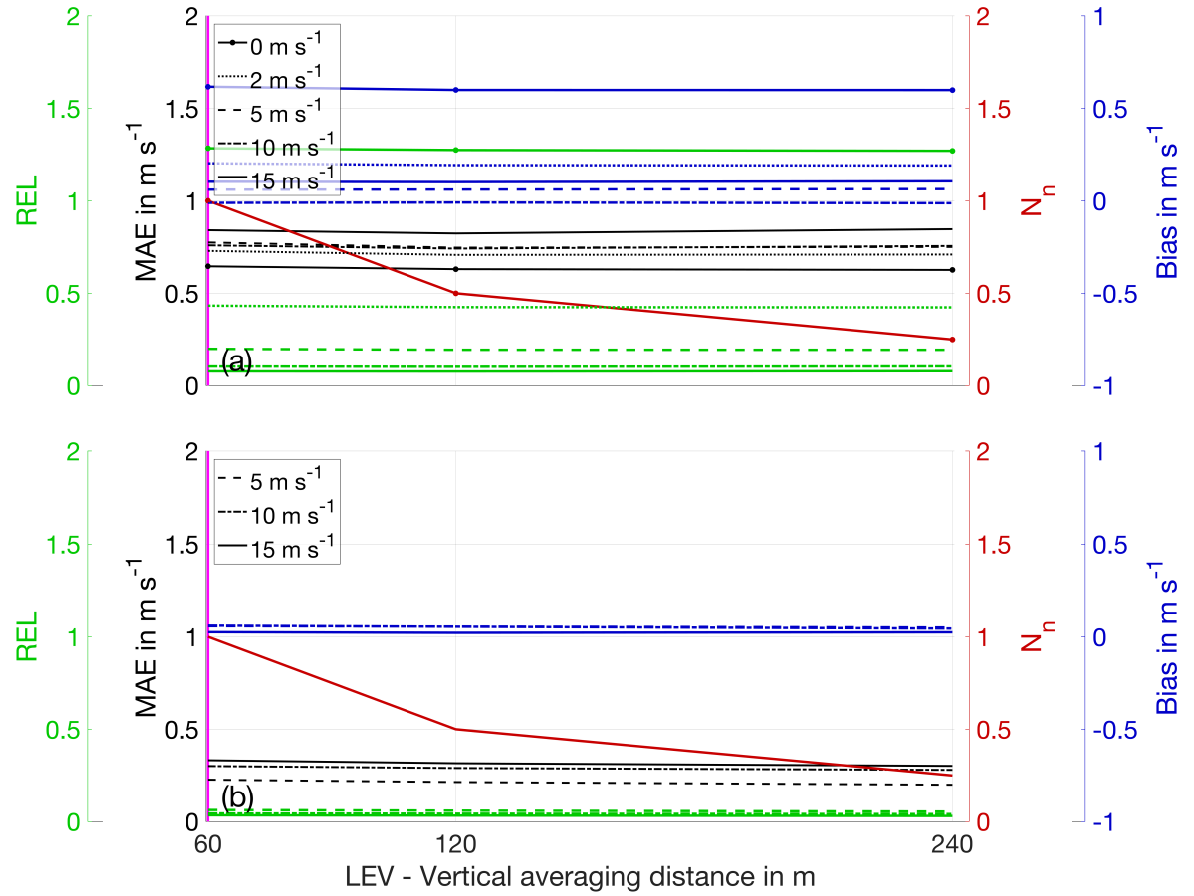

Fig. 5.9.: REL, MAE, N_n and bias for three vertical averaging distances using STP and RET otherwise. a) LES set A with high surface sensible heat flux and five background wind cases. b) LES set B with low surface sensible heat flux and three background wind cases. The vertical averaging distance of the standard retrieval strategy is indicated as a magenta line.

Another parameter which can be varied is the wind profile resolution in the vertical direction through the vertical averaging distance (LEV). Depending on the vertical retrieval volume extent the number of range gates used to retrieve one wind profile point differs and can influence the LSQ-fit quality. The vertical averaging distance is varied between 60 m, 120 m and 240 m, which corresponds to having one, two or four range gates in the retrieval volume at the respective altitudes for 60° scan elevation. The influence of varying the vertical averaging distance is hardly discernible in terms of wind profile error (Fig. 5.9). However, the number of retrieved wind profile points varies strongly because the vertical profile resolution is changed. The reason for vertical averaging having no influence is founded in the turbulence structure (similar to the indifference to measurement frequency changes). 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 248 m vertical averaging distance, but at the cost of having far fewer profile points available. Decreased vertical resolution is generally undesirable and can be problematic in regions with vertical shear of the horizontal wind. 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 for single profiles, but as this error averages to zero over many profiles it is more accurately described as a random error, see Fig. 4.3 for an illustration). This correlation is the reason why the error discussed here is impossible to identify visually in real Doppler lidar wind profiles: No small-scale variations occur, but rather offsets over large sections of the profile. The magnitude of the random error only becomes noticeable when looking at profiles sampled at different locations or times. The disguise discussed here is likely one of the reasons why the error due to flow inhomogeneity has received little attention in literature so far. The vertical correlation effect is present in both the ADLS and real measurements with equal magnitude.

Sector blanking

Last, the robustness of the retrieved wind profiles to partial data unavailability (BLA), termed 'blanking', is investigated (Fig. 5.10). In real-world ADL measurements, sectors of the AVAD scan can be unavailable due to the presence of clouds, obstacles such as mountains or other disturbances. The effects are simulated by discarding a specific sector of the scan geometry before wind profile retrieval (in the ground-based reference system ENU, in which the disturbance would be present). For analysis, five cases are considered: sector blanking from $0-90°$, $0-135°$, $0-180°$, $0-225°$ and $0-270°$. A sector blanking of 90° leads to a moderate increase in the average error levels for all background wind cases, without loss in the number of retrieved wind profile points. With a further increase to 180° sector blanking the wind profiling quality is more strongly degraded and a noticeable bias is introduced. Nevertheless, the full number of wind profile points remain retrievable. Beyond 180° sector blanking severe wind profiling bias develops also in the low turbulence case and profiling error grows rapidly. With 270° sector blanking, the minimum CN is above 20 and relaxed CN quality control criteria ($CN < 25$) have to be applied to enable retrieval of any wind profiles. Consequently, due to the relaxed quality control criteria, no reliable wind profile retrieval is achievable anymore.

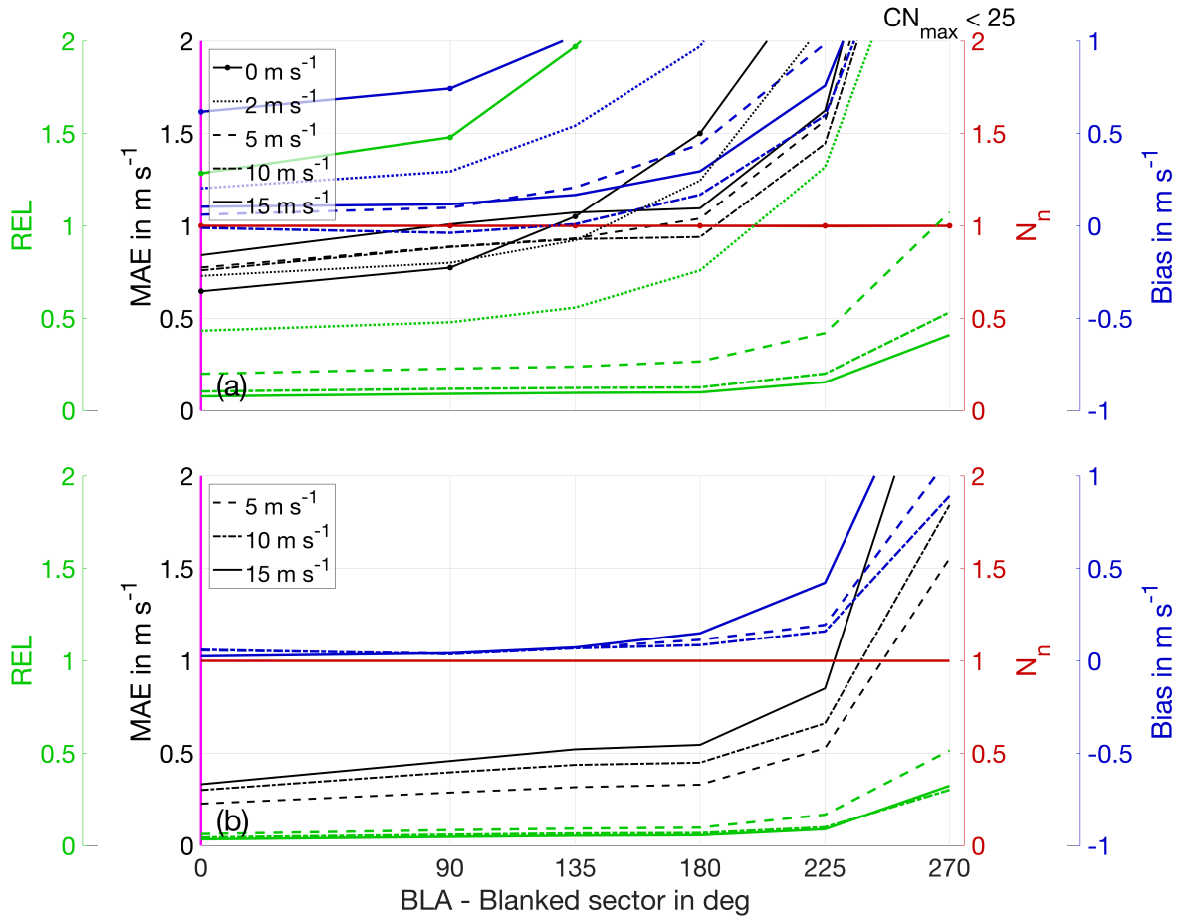

Fig. 5.10.: REL, MAE, N_n and bias for four sector blankings using STP and RET otherwise. a) LES set A with high surface sensible heat flux and five background wind cases. b) LES set B with low surface sensible heat flux and three background wind cases. The sector blanking of the standard retrieval strategy is indicated as a magenta line. For BLA = 270° the *CN* quality control criteria is relaxed to $CN < 25$ as no retrievals fulfill the standard $CN < 12$ threshold.

An optimized system setup and retrieval strategy

Concluding, in inhomogeneous flow conditions shallow scan elevations (e.g. 45°) should be utilized to limit the impact of the vertical wind on retrieval quality. The increase in across-track measurement footprint can be offset by keeping the retrieval volume confined in along-track direction using a VVP technique. This is possible, because ADLS results show that forward and aft measurements by the lidar are sufficient to allow for reliable retrieval of the wind vector without utilizing the full retrieval volume, also for temporally separated measurements from different aircraft positions. Sufficient exploration of the retrieval volume can be ensured through the reliable *CN*-based quality control. Thus, high spatial resolution with a confined measurement footprint can be achieved, allowing for high-resolution sampling of flow phenomena in along-track direction (in the limit, a dual-Doppler technique with additional retrieval of the

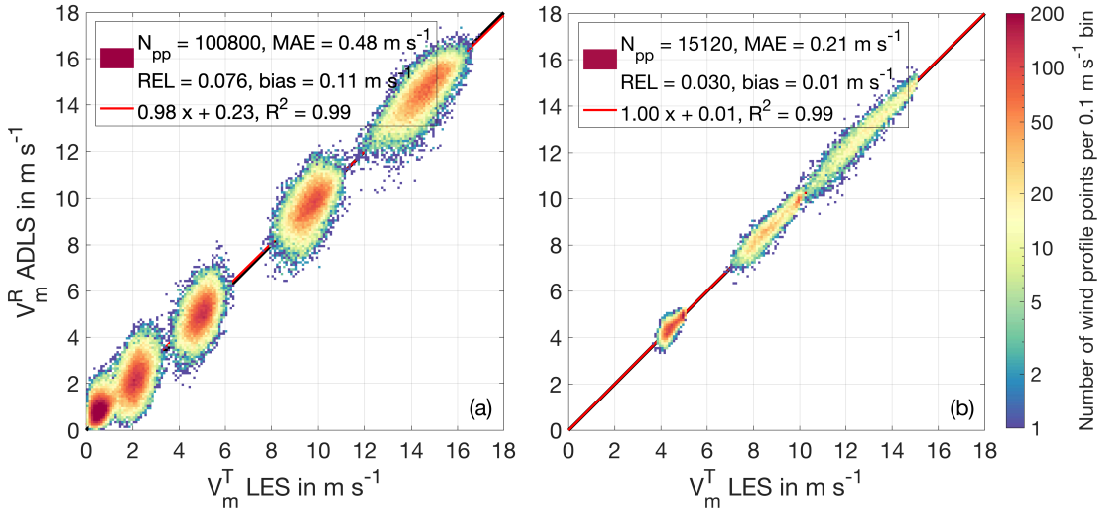

Fig. 5.11.: Comparison of LES truth and AVAD retrieved wind speed for an optimized measurement system and retrieval strategy for (a) LES set A and (b) LES set B. Color-coded is the histogram distribution of all measurements. The observable deviations from the 1:1 line reveal the AVAD retrieval error due to the violation of the flow homogeneity assumption.

across-track wind component is reached). This approach is vastly preferable to the second option of choosing a steep scan elevation angle (limited across-track footprint) and increased along-track averaging distance, because many more wind profiles can be retrieved using the former approach (besides the potential retrieval biases present in the latter approach).

Alongside the shallow scan elevation angle, a system with high scan speeds allows for improved retrieval quality and increased along-track resolution. In the case of high scan speeds, high measurement frequencies are preferable and can provide added value (because the correlation between the measurements is reduced by the fast scan speeds). Vertically, the retrieval may utilize the highest resolution possible, but one should be aware of the disguised, vertically correlated wind profiling error due to vertically correlated turbulence. Sector blanking $< 180°$ is unproblematic, its impact can be reliably controlled through the application of CN-based quality control.

Based on these insights, an optimized system setup and retrieval strategy for ADL wind profiling inside the BL at highest resolution can be presented. To do so, a simulation of a system scanning with 45° elevation, $45\,°\,s^{-1}$ scan speed and 10 Hz measurement frequency is simulated. The associated retrieval is conducted with an along-track averaging distance of only 325 m and increased across-track averaging distance of 2200 m. Thereby, the retrieval footprint is reduced by more than 30% from the standard retrieval strategy.

The resulting wind profile retrieval characteristics are shown in Fig. 5.11. The wind profiling error is strongly reduced by approximately 30% (both smaller MAE and bias) for both atmo-

spheric conditions, despite the smaller measurement footprint and increased retrieval resolution, compared to the standard system setup and retrieval strategy (Fig. 5.1). The evolution of the wind field between temporally separated forward and aft observation of the ADL into the small retrieval volumes does not impede retrieval accuracy, as is shown by the results from LES set B which considers the time evolving wind field. Further, due to the short horizontal averaging distance, four times more wind profiles can be retrieved, which is an invaluable asset in real-world measurements. The increased vertical resolution due to the more shallow scan also results in the retrieval of more wind profile points for every profile. Therefore, an increase by a factor of five in the number of retrieved wind profile points is obtained overall.
The optimized system setup and retrieval strategy allow for investigation of flow phenomena with one symmetry axis at highest resolution (e.g. when the aircraft is able to fly perpendicular to the flow phenomena). Examples for such flow phenomena with symmetry in the across-track direction are e.g. flow across a mountain range, wind turbine wakes and cold-pool outflows. Of course, the improvement of measurement strategies for other flow phenomena is also possible using the ADLS (e.g. a focus on higher retrieval accuracy at degraded profile resolution). Thus, the ADLS underlines its effectiveness and utility in the analysis and design of measurement strategies for ADL BL measurements.

5.4. Evaluation of the data-driven uncertainty estimation method

The previous sections show that inhomogeneous flow conditions present an important source of wind profiling error which cannot be fully mitigated, even when using the most adequate system setups and retrieval strategies. Therefore, the possibility of a data-driven uncertainty estimation based on the measured data is investigated here. A data-driven uncertainty estimation method as proposed in Sec. 3.6 offers the possibility to assess the reliability of the retrieved wind profiles based on the measured data alone. Thereby, uncertainty estimates can be assigned individually to every retrieved wind profile point, without the need to know the specific atmospheric conditions in the measurement volume or even the input truth (which is generally unavailable in real-world measurements).
The ADLS offers the possibility to evaluate any uncertainty estimation method for its validity and applicability in terms of both absolute and statistical values. Both the retrieved wind profile as well as the input wind profile are known exactly, thereby the retrieval error is directly accessible as complications such as the representation error can be avoided. Therefore, the ADLS implementation of the theory outlined in Sec. 3.6 allows for an investigation whether a data-driven uncertainty estimation is possible and reliable. In this section the metrics used

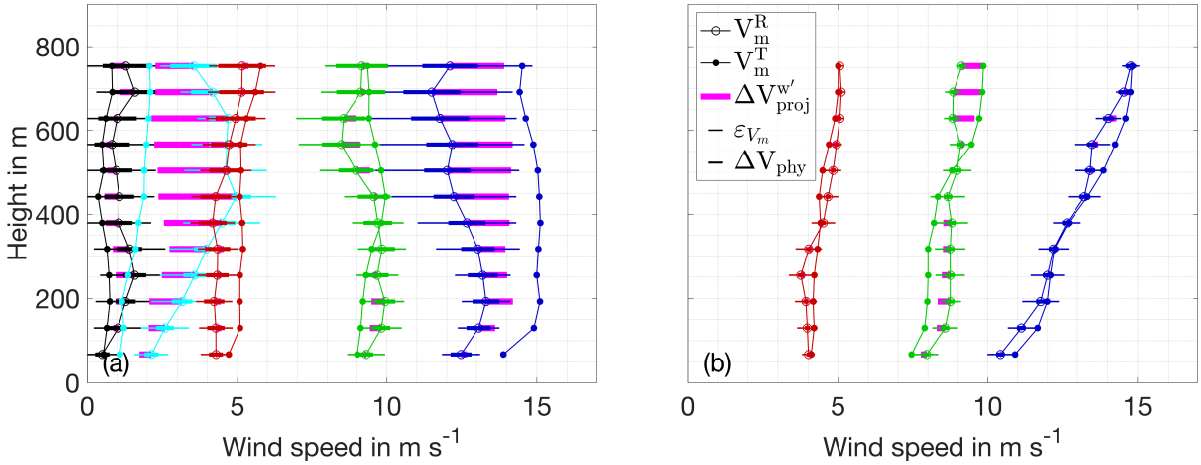

Fig. 5.12.: Examples of retrieved wind profiles and the associated sub- and supra-scan-volume uncertainties for (a) LES set A and (b) LES set B using an ideal measurement system with STP and RET. $\Delta V_{proj}^{w'}$ quantifies the specific supra-scan-volume contribution present but is only available if the input truth is known. The other uncertainties are also available in real-world measurements and represent the data-driven uncertainty estimation. The underlying LES input truth, simulated radial velocities and the LSQ-fit are shown in Fig. 4.4 for the 320 m altitude level.

for evaluation are introduced and the investigation is conducted for the standard system setup and retrieval strategy as an example. In Sec. 5.4.2, the analysis is extended to all system setups and retrieval strategies discussed previously, to prove the validity and applicability of the uncertainty estimation method for a wide range of system setups and retrieval strategies.

Figure 5.12 displays examples of the uncertainty estimation obtained using the theory outlined in Sec. 3.6 and Appendix A.3, for the same wind profiles as shown in Fig. 4.3. The profiles illustrate the added value of the data-driven uncertainty estimates developed as a part of this work. Previously, no data-driven uncertainty estimates could be assigned to the retrieved wind profiles, which is clearly problematic given the large observed errors due to flow inhomogeneity. At first glimpse, the magnitude of the data-driven uncertainty estimates appear related to the magnitude of error present, a promising sign which is thoroughly checked in the following sections. The data-driven uncertainty estimation exhibits varying levels and ratios of resolved sub-scan-volume uncertainty and parameterized supra-scan-volume uncertainty. Additionally, both contributions vary with altitude due to the changing sampling as well as turbulence characteristics. As expected, the exact amount of supra-scan-volume flow inhomogeneity and the associated erroneous mapping can only be determined and corrected with full knowledge of the wind field in the ADLS. This is especially noticeable for retrievals with large supra-scan-volume contributions, for example the profile retrieved from the 2 and 15 $\mathrm{m\,s^{-1}}$ background wind cases for LES set A. Another prominent example occurs in the entrainment region of the

profile retrieved from the $10\,\mathrm{m\,s^{-1}}$ background wind case for LES set B. There, gravity waves are present which lead to non-negligible supra-scan-volume contributions. In this region, the parameterized supra-scan-volume uncertainty is not expected to capture the full extent of supra-scan-volume error for individual profiles. Nevertheless, the parameterized supra-scan-volume uncertainty should capture the average magnitude of erroneous mapping, which is checked in the following.

In order to evaluate the validity and applicability of the data-driven uncertainty estimation method, two characteristics are checked. First, the mean estimated uncertainty (UNC) should on average be equal to or larger than the ADLS observed MAE for a conservative estimation. Second, the distribution of the observed error compared to the estimated uncertainty should follow a Gaussian distribution with $\sigma_G < 1$ for a conservative estimation. In this case, at least 63% of the ADLS observed true wind profile points lie within $\pm 1\sigma_G$ and 95.5% lie within $\pm 2\sigma_G$ of the estimated uncertainty, as is expected for a conservative estimation. A $\sigma_G > 1$ corresponds to less than 63% of the observed wind profile points falling within $\pm 1\sigma_G$, representing a non-desirable underestimation of uncertainty from a statistical perspective. The latter case can occur even when the UNC is larger than the observed MAE, which makes the check of both absolute and statistical properties necessary. The calculation of both metrics is detailed in the following.

On average, the UNC should be larger than the ADLS observed MAE. The UNC consists of two contributions, the sub-scan-volume uncertainty and the supra-scan-volume uncertainty (Sec. 3.6). The sub-scan-volume uncertainty contribution is calculated using the covariance matrix based uncertainty propagation and utilizes the effective sample-size corrected residual velocity variance as input. The supra-scan-volume contribution can be calculated using two different approaches (Sec. 3.6 and Appendix A.3). If the supra-scan-volume contribution is determined directly from the ADLS, the overall UNC is calculated in the following way:

$$\mathrm{UNC}_{dir} = \Sigma_n^N \left(\varepsilon_{V_{m\,n}} + |\Delta \mathrm{V}^{w'}_{proj\,n}| \right) / N. \tag{5.4}$$

The following discussion also uses subsets of these uncertainty contributions to explain the functionality. These are

1. UNC_{cov} for the raw covariance matrix based uncertainty estimate with neither effective sample size correction nor supra-scan-volume correction,

2. UNC_{eff} for the covariance matrix based uncertainty estimate with only effective sample size correction (no supra-scan-volume correction) and

3. UNC_{sup} for the uncertainty estimate with only supra-scan-volume correction (no covariance matrix based sub-scan-volume correction).

Correspondingly, if the supra-scan-volume contribution is parameterized statistically based on physical considerations (Appendix A.3) the UNC is calculated using

$$\mathrm{UNC}_{phy} = \Sigma_n^N \left(\varepsilon_{V_{m\ n}} + \Delta V_{phy\ n}\right)/N. \tag{5.5}$$

Besides the average values of the uncertainty estimates the statistical properties also have to be evaluated. In this work, the evaluation is done with the use of histograms relating the ADLS observed wind profile error to the estimated uncertainty. The individual location of a histogram member is given by

$$p_{his\ n} = \frac{(V^R_{m\ n} + \Delta \mathrm{V}^{w'}_{n\ proj}) - V^T_{m\ n}}{\varepsilon_{V_{m\ n}}} \tag{5.6}$$

for the direct approach and correspondingly for the subsets discussed above. For the parameterized approach the individual histogram member location is given by

$$p_{his\ n} = \frac{V^R_{m\ n} - V^T_{m\ n}}{\varepsilon_{V_{m\ n}} + \Delta V_{phy\ n}}, \tag{5.7}$$

because the sign of the supra-scan-volume contribution is unknown. After obtaining the histogram from all members, Gaussian distributions are fitted to the histograms using a least-squares minimization.

5.4.1. Functionality for standard system setup and retrieval strategy

This section investigates the question whether the uncertainty of the retrieved profile can be estimated based on the measurements from a statistical point of view. Statistical results of various uncertainty estimation methods are shown in Fig. 5.13 for LES set A using STP and RET for three measurement frequencies FME = $1, 5, 10$ Hz.

A number of general observations can be made from the distributions which are present in all of the panels. The uncertainty estimate based on direct estimation through the covariance matrix (UNC_{cov}), using the residual variance directly as a proxy for data variance, does not have predictive quality, especially for higher measurement frequencies. Two reasons are responsible for this uncertainty estimation failure. Firstly, the data variance is greatly underestimated due to strong correlation among the residuals in the LSQ-fit. Using the effective sample size estimation

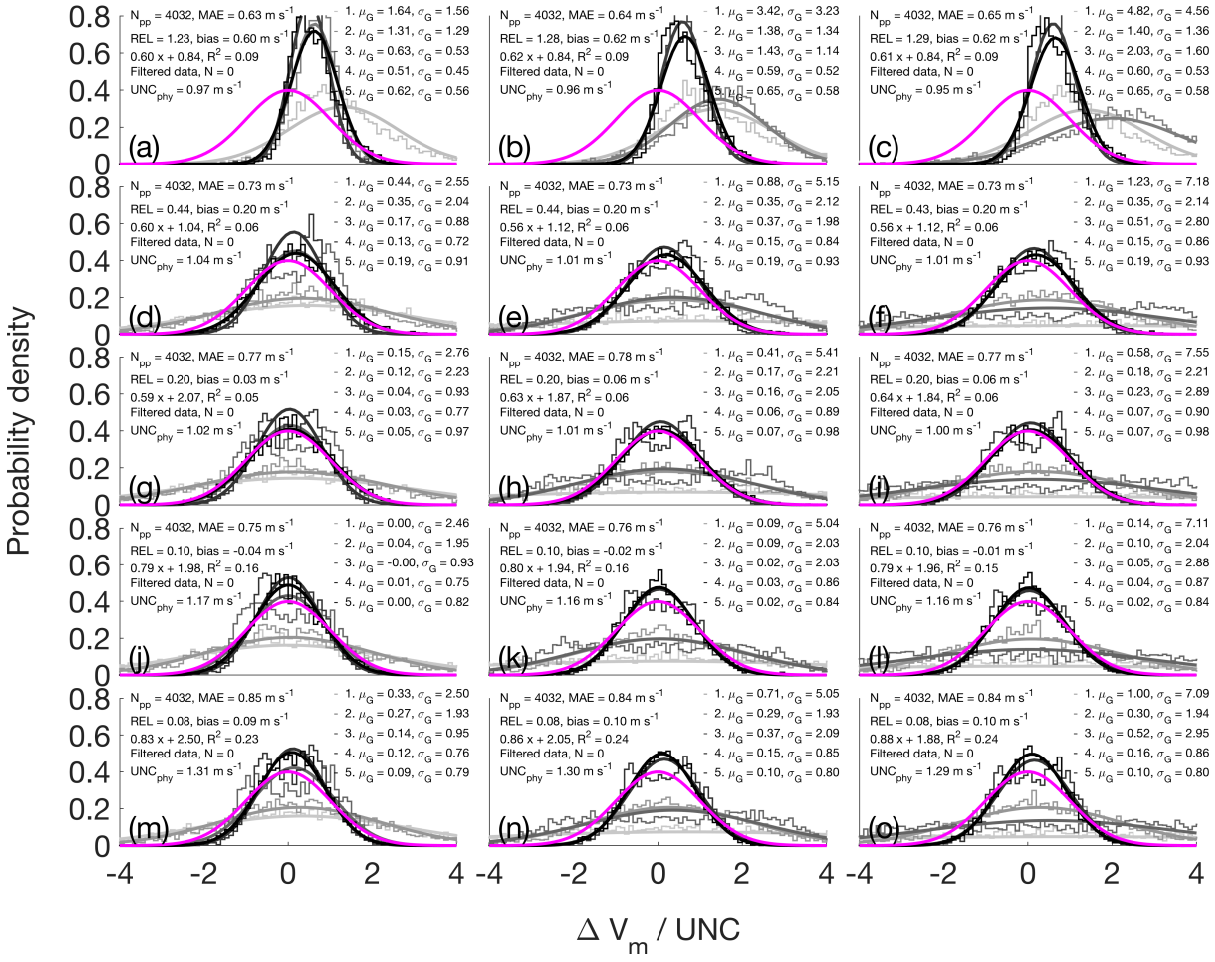

Fig. 5.13.: Statistical properties of different uncertainty estimation methods for an ideal measurement system with STP and RET using LES set A.
(a), (b), (c) 0 $\mathrm{m\,s^{-1}}$ background wind case, FME = 1, 5, 10 Hz respectively.
(d), (e), (f) 2 $\mathrm{m\,s^{-1}}$ background wind case, FME = 1, 5, 10 Hz respectively.
(g), (h), (i) 5 $\mathrm{m\,s^{-1}}$ background wind case, FME = 1, 5, 10 Hz respectively.
(j), (k), (l) 10 $\mathrm{m\,s^{-1}}$ background wind case, FME = 1, 5, 10 Hz respectively.
(m), (n), (o) 15 $\mathrm{m\,s^{-1}}$ background wind case, FME = 1, 5, 10 Hz respectively.
The histograms displayed show the distributions of the following uncertainty estimation methods:
1. 20 % opaque, UNC_{cov}, uncorrected residual variance propagation through covariance matrix.
2. 40 % opaque, UNC_{eff}, only including effective sample size correction, taking into account correlation in measurements due to high measurement frequency.
3. 60 % opaque, UNC_{sup}, only including ADLS-based correction for supra-scan-volume contributions.
4. 80 % opaque, UNC_{dir}, effective sample size correction and ADLS based supra-scan-volume contribution.
5. 100 % opaque, UNC_{phy}, effective sample size correction and physically parameterized supra-scan-volume contribution (data-driven real-world method).
Gaussian fits to every histogram are displayed with the same color. The magenta line is the envelope of a Gaussian distribution with $\mu = 0$, $\sigma = 1$.

increases the quality of the uncertainty estimate noticeably (UNC_{eff}), but still does not yield the desired Gaussian distribution. The second problem is the presence of supra-scan-volume flow inhomogeneity in the wind field, which is not detectable in the LSQ-fit residuals. The effect of supra-scan-volume flow inhomogeneity can be analyzed and taken into account using the ADLS, but by itself it also does not lead to a reliable uncertainty estimation (UNC_{sup}).
Due to the biased retrieval at low wind speeds the distributions exhibit mean values shifted slightly towards the positive (Fig. 5.13 a, b, c). UNC_{cov} shows a higher μ_G than UNC_{eff}, even though they originate from the same wind profile points without bias correction. This difference is because the bias is more influential if the estimated uncertainty is small, as in this case the members of the histogram are located more towards positive values (consequently μ_G is higher). The uncertainty estimation is improved and yields the desired conservative estimation and Gaussian distribution qualities, if both the effective sample size corrected sub-scan-volume residual variance uncertainty propagation through the covariance matrix and the ADLS based supra-scan-volume flow inhomogeneity error analysis are combined (UNC_{dir}). Results are independent from background wind speed (except for influences by the retrieval bias), proving the general suitability of the covariance based uncertainty estimation approach using the residual velocity variance as a proxy for data variance. Further, the results are also independent of the measurement frequency, proving the suitability of the effective sample size estimation to account for correlation among the measurements (Eq. 3.41). It should be noted that the model based supra-scan-volume analysis does not fully correct the retrieval bias at low wind speeds, hence, a $\mu_G > 0$ remains. The skewed distribution is because the sub-scan-volume contribution is not corrected by the direct approach, thereby, members of the histograms cluster on the positive side. Similarly, for higher wind speeds two bumps in the vicinity of $+1\sigma_G$ and $-1\sigma_G$ are detectable. These bumps are also caused by the partial correction due to the exact ADLS based supra-scan-volume analysis, systematically leaving $\pm 1\sigma_G$ of the sub-scan-volume uncertainty analysis for correction.
The data-driven uncertainty estimation (UNC_{phy}), based on a physically parameterized supra-scan-volume error contribution, also presents a successful conservative uncertainty estimation for all cases. The distribution and fitted Gaussian curves of UNC_{dir} and UNC_{phy} are very similar for both σ_G and μ_G, showing the good quality of the physical parameterization. It should be noted that the data-driven uncertainty estimate cannot correct the retrieval bias at low wind speeds, because it does not yield signed supra-scan-volume contributions but rather average, absolute values of uncertainty that can be positive or negative. Hence, μ_G is a function of the retrieval bias and the statistical distributions do not posses a μ_G close to zero for low wind speeds. Nevertheless, the data-driven uncertainty estimation does capture the average

magnitude of supra-scan-volume flow inhomogeneity, as shown by the σ_G values below one for the 0 and 2 $\mathrm{m\,s^{-1}}$ background wind cases.

The following analysis on the applicability and reliability of the uncertainty estimation method for different system setups and retrieval strategies is limited to the 10 Hz measurement frequency case. The additional problems due to correlation of neighboring radial velocity measurements make an estimation of the effective sample size necessary. Thus, the 10 Hz measurement frequency case is the most challenging uncertainty estimation to perform. Though the 1 Hz uncertainty estimates are not shown in the following, its results are always slightly more conservative than those of the 10 Hz case. Further, the analysis is limited to an evaluation of the data-driven UNC_{phy} uncertainty estimation, in order to limit the length of discussion and because it is the one applicable to real-world measurements. UNC_{phy} also presents the most challenging method as no information from the LES is introduced except for the BL height and the vertical profile of the mean absolute vertical wind (which can both be extracted from nadir measurements in real-world scenarios). Therefore, the condensed results are presented by displaying only σ_G and μ_G of the data-driven UNC_{phy} uncertainty estimation histograms.

Vertical profile of uncertainty estimation parameters

The two LES sets A and B exhibit differing vertical distributions of wind profiling error, as well as different contributions of sub- vs. supra-scan-volume flow inhomogeneity due to the different BL heights. Further, the scan-volume diameter also varies with altitude, changing the ratio of resolved sub-scan-volume and parameterized supra-scan-volume uncertainty contributions. The vertical changes in profiling error characteristics have to be captured and resolved by a reliable data-driven uncertainty estimation method.

Figure 5.14 shows the vertical profile of the condensed uncertainty estimation quality evaluation parameters for the most challenging estimation method UNC_{phy}, with the aircraft flying at the standard flight altitude of 1100 m. Figure C.10 shows the vertical profile with the aircraft flying at 1700 m, changing the ratio between the sub- and supra-scan-volume contributions and exploring the full BL for both LES sets.

The increase in wind profiling MAE from the surface towards the middle of the BL (where up- and downdrafts have maximum intensity) is captured by the UNC_{phy} uncertainty estimate, as is the decrease towards the BL top for both flight altitudes (Fig. 5.14, C.10). Further, the differences in error magnitude between the background wind cases are also represented in the uncertainty estimation (e.g. 15 $\mathrm{m\,s^{-1}}$ having the largest profiling error). The uncertainty estimation is conservative at all levels inside the BL, UNC_{phy} overestimates the observed MAE

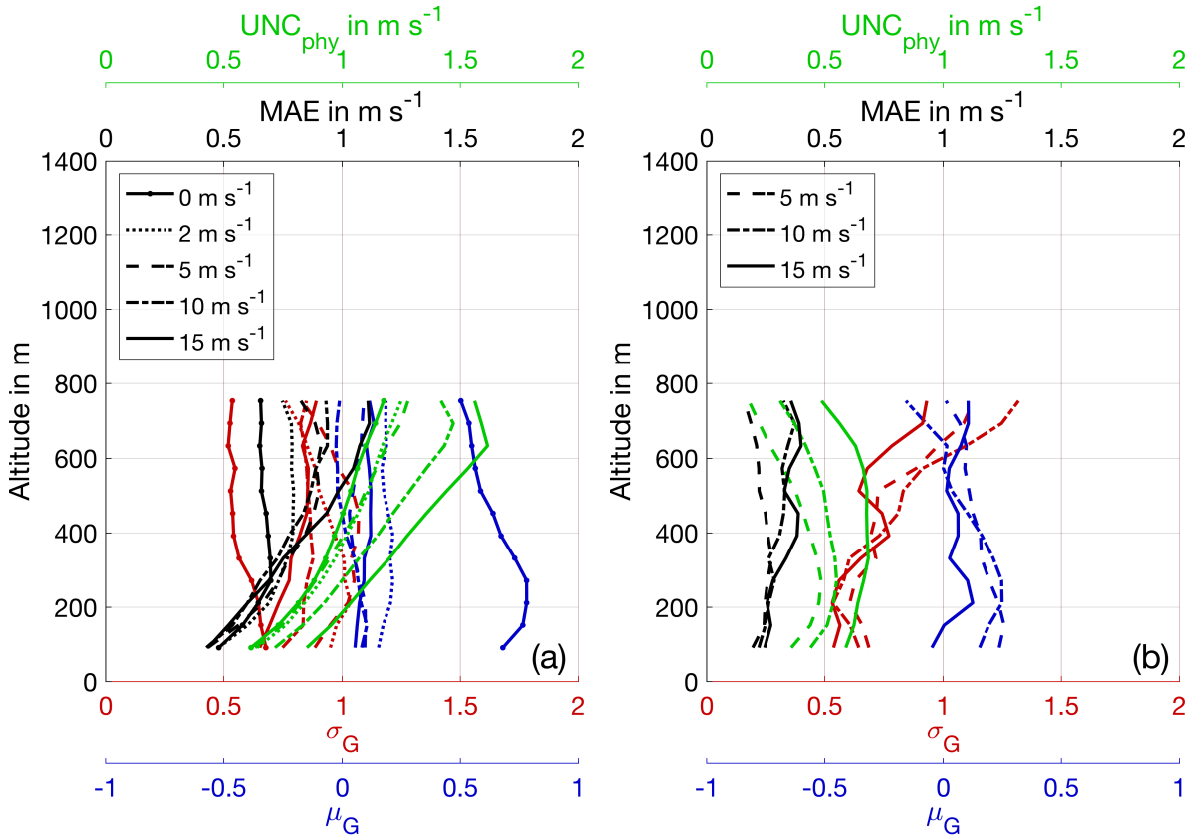

Fig. 5.14.: Profile of uncertainty estimation parameters UNC_{phy}, MAE, μ_G and σ_G as a function of height. Displayed data are for an ideal system using the standard system setup (aircraft flying at 1100 m) and retrieval strategy for (a) LES set A and (b) LES set B.

by approximately a factor of 2. Only small variations in σ_G are observed despite the changing error magnitude. Due to the positive retrieval bias μ_G is not centered on 0 but shows slightly positive values up to 0.7 instead for LES set A and B, especially for the $0\,\mathrm{m\,s^{-1}}$ background wind case. The uncertainty estimation becomes more conservative with decreasing wind speed. Potentially, this is due to the decreased variance of the low background wind case retrievals (because they are confined to positive wind speeds), which is not accounted for by the uncertainty estimation method. Elevated σ_G levels are present in the entrainment zone and above for LES set B, although the average UNC_{phy} remains larger than the MAE (as well as generally more variation due to the smaller sample size compared to LES set A, Fig. 5.14b). The slight statistical underestimation is due to the supra-scan-volume parameterization, which is designed for BL turbulence. The noticeable presence of gravity waves in the entrainment zone of LES set B (see Fig. 4.2 for their vertical wind magnitude, Fig. 5.12 for the resulting retrieval error and uncertainty estimation failure) violates the assumptions of the parameterization for the low flight altitude (e.g on the dominant eddy size). When the aircraft is flying higher, the gravity waves in the entrainment zone are captured through the resolved sub-scan-volume uncertainty

contribution (because the scan-volume is increased), consequently the uncertainty estimation is improved (Fig. C.10b). Overall, it should be noted that the data-driven uncertainty estimation method is far from being unreliable even in this region and presents a strong improvement compared to the more basic approaches discussed above. Future work can improve the parameterization scheme to also include the effect of gravity waves.

5.4.2. Reliability and applicability

The previous sections show the possibility of a data-driven uncertainty estimation for the standard system setup and retrieval strategy. The wind profiling error magnitude and characteristics change for other system setups and retrieval strategies, as do the contributions and the ratio of resolved sub-scan-volume and parameterized supra-scan-volume uncertainty contributions. Consequently, the appropriate functionality of the data-driven uncertainty estimation method needs to be validated for these cases as well, which is done in the next section. To do so, the reliability of the uncertainty estimation is evaluated for the same range of system setups and retrieval strategies as in Sec. 5.2 and 5.3.

The reliability of the data-driven uncertainty estimation for varying scan elevation angles can be evaluated from Fig. 5.15. The ADLS results (see also Fig. 5.5) show a strongly increasing MAE with steeper scan elevation angle, due to the increased impact of the vertical wind contribution as well as a biased retrieval discussed previously. The estimated UNC_{phy} increases in-line with the increased retrieval error up to the highest levels with $\mathrm{MAE} > 2\,\mathrm{m\,s^{-1}}$, as desired. On average, the estimated UNC_{phy} values are larger than the observed MAE for all atmospheric conditions, background wind cases and scan elevations, yielding a conservative estimation. The different error levels of different atmospheric conditions and background wind cases are also captured by the uncertainty estimation. For example, the uncertainty estimated for LES set A is much higher than that estimated for LES set B. Similarly, the uncertainties estimated for higher background wind cases are higher than those for lower background wind cases, in-line with the observed MAE. The estimation is slightly more conservative for steep scan elevation angles, due to the stronger influence of the vertical wind on the radial velocity variance. With less influence of the vertical wind for shallow scan elevation angles, the overestimation is reduced and σ_G approaches 1. A small excess of $\sigma_G > 1$ is observed for LES set A at 30° scan elevation. This excess is because the supra-scan-volume parameterization is based on w and does not consider flow inhomogeneity of the u or v component which can also influence retrieval error, especially for shallow scan elevations. The small σ_G excess is unproblematic as the UNC_{phy} estimation still exceeds the MAE level (and errors are generally smaller). As before, the μ_G of

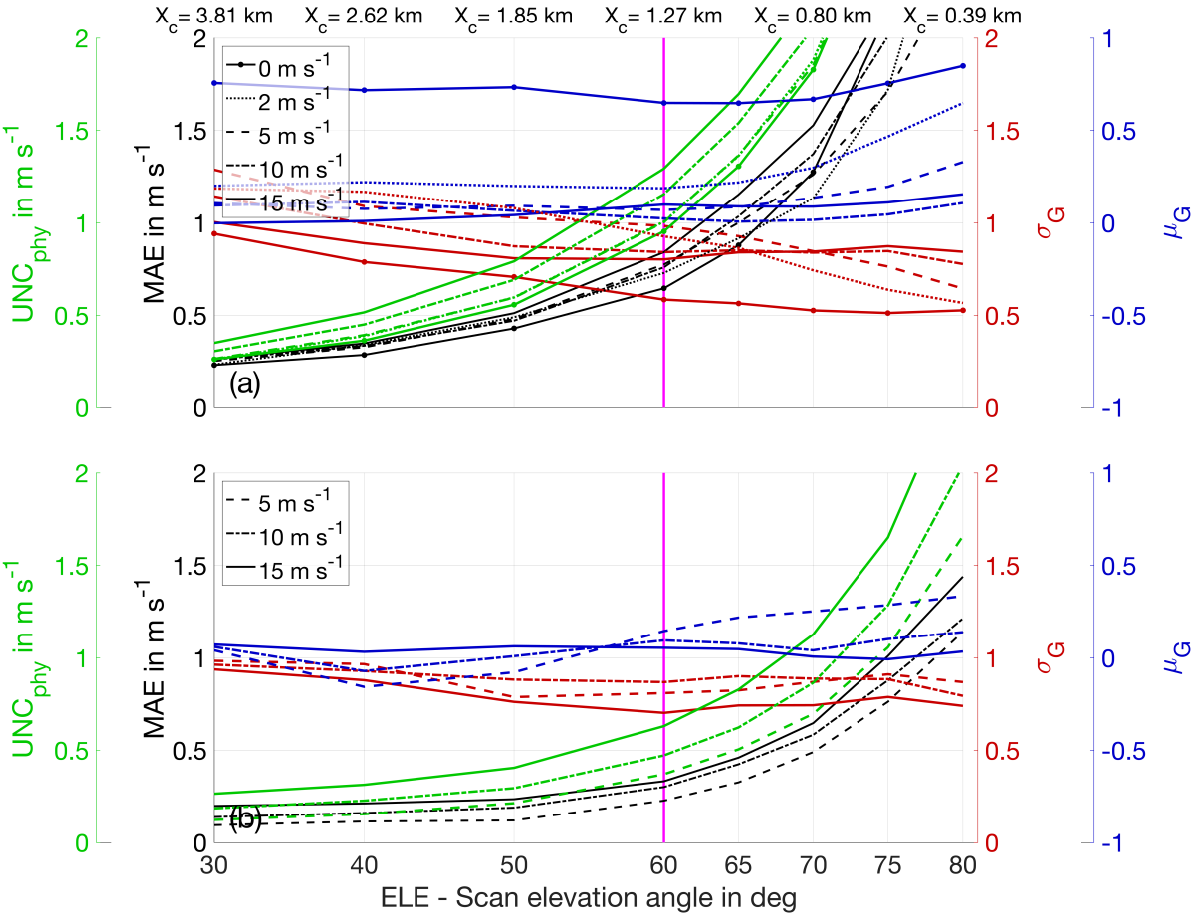

Fig. 5.15.: Data-driven uncertainty estimation quality for six scan elevation angles and STP and RET otherwise. (a) UNC_{phy}, MAE, μ_G and σ_G obtained from the histogram analysis for LES set A with high surface sensible heat flux and (b) for LES set B with low surface sensible heat flux. The scan elevation angle of the standard system setup is indicated as a magenta line.

the distributions fluctuate around 0 for all background wind cases, except for cases when the underlying retrieval is biased.

The reliability of the data-driven uncertainty estimation method for varying scan speeds can be evaluated from Fig. C.11. A decrease of profiling error was observed for faster scans, due to better explored retrieval volumes and reduced correlation between the measurements. The decrease is also present in the estimated uncertainty, although at reduced magnitude. For faster scans the uncertainty estimation becomes more conservative because the parameterized supra-scan-volume parameterization is of the same magnitude for all scan speeds. This constant level leads to an overestimation, because the better exploration of the scan volume does identify non-perfect horizontal gradients of the vertical wind better. Thereby, the magnitude of erroneous mapping is reduced and flow inhomogeneity is resolved in the residuals. The better exploration of the volume can be accounted for through improved parameterization in future work (e.g

by multiplying with a factor accounting for the number of rotations). A weak uptick of σ_G is observed at the fastest scan speed of $90\,^\circ\,\mathrm{s}^{-1}$ (shortest scan duration of 4 s). This uptick is because measurement positions from subsequent scan rotations are separated by only $4\,\mathrm{s} \cdot 65\,\mathrm{m\,s}^{-1} = 260\,\mathrm{m}$ in along-track distance for the fastest scan speed. This distance is below the integral length scale of turbulence encountered in the LES, therefore the measurements are not completely independent but correlated instead. This correlation is not accounted for by the effective sample size correction, which only takes into account the correlation of measurements in the temporal domain. Consequently, the effective sample size reduction is not captured fully anymore, which is unproblematic given the general conservative overestimation.

The retrieval strategy also influences measurement quality and thereby the data-driven uncertainty estimation method, which also needs to be validated hence. The reliability of the data-driven uncertainty estimation for along-track averaging distances can be evaluated from Fig. C.12. Generally, the increase in error levels with shorter averaging distances is captured, as well as the decrease for longer averaging distances. σ_G varies only slightly, showing the reliability of the data-driven uncertainty estimation also for challenging conditions with application of CN-based quality control.

Increased vertical averaging does not improve retrieval quality noticeably due to the strong vertical correlation of turbulence. Hence, the uncertainty estimation should also be unaffected, despite the increased number of measurement available. The independent estimation of the uncertainty method can be validated from Fig. C.13, proving the suitability of the along-beam sample size reduction formula (Eq. 3.40).

Last, the reliability of the data-driven uncertainty estimation method under the presence of sector blanking can be evaluated from Fig. C.14. The estimated UNC_{phy} captures the increase in error levels up to a sector blanking of 180° conservatively. Above this, while the estimated uncertainty captures the strong increase in MAE and yields a conservative estimation of the mean error level, there is a slight increase of σ_G to levels greater one for a sector blanking $> 225°$ (CN-based quality control was relaxed to $CN_{max} < 25$ in this region). The slight statistical underestimation of uncertainty is due to difficulties in estimating the effective sample size reliably for very short measurement segments (and thereby very few data points) and because of increasing retrieval bias. Although UNC_{phy} levels show a conservative estimation on average, the effective sample size is not always estimated reliably, which can lead to elevated tails of the Gaussian distribution. The increasing retrieval bias is also the reason for the slight underestimation of the average MAE level for the $0\,\mathrm{m\,s}^{-1}$ background wind case for very large sector blankings. Hence, the uncertainty estimation method is best used together with the application of reasonable CN quality control criteria, e.g. $CN_{max} < 12$. Overall, even in this case, the

data-driven uncertainty estimation method presents a one order of magnitude better uncertainty estimation compared to the direct covariance matrix based approach. Further, the absolute error levels of the uncertainty estimation qualify the retrieval as unusable, thereby the statistical properties are not as relevant as the overall error level.

6. Assessment and application of airborne Doppler lidar wind observations using a prototype system

In this chapter, results obtained with a new prototype ADL system based on four research flights conducted in the vicinity of Brunswick (Germany) are presented. It is demonstrated that the new measurement system, onboard a low and slow flying medium-range turboprop aircraft, enables high resolution and high accuracy measurements of the turbulent BL wind field. The uncertainties associated with real-world ADL measurements inside the BL are analyzed and quantified, both for wind profiling and nadir retrievals. To this end, existing procedures from a number of previous studies are combined, refined and extended, leading to an in-depth data processing and quality controlling scheme. Thus, it is shown that the new system enables the detailed study of turbulence characteristics in the BL. The usefulness of such ADL observations is demonstrated through the analysis of three application studies, which show that the ADL can provide substantial added value.

6.1. Prototype system and measurement flights overview

In order to assess the potential of ADL wind measurements for BL research, four research test flights were conducted in the vicinity of Brunswick in summer 2019. The characteristics of the aircraft and lidar system used for the research flights are described in Sec. 2.3.3. Unfortunately, due to unforeseen difficulties in the development process as well as management problems, the intended scanner system could not be readied by the manufacturing company in time. Consequently, the scanner was not available during the research flights. Therefore, a prototype version of the system was designed and certified for flight as a part of this work. The prototype system does not utilize a scanner, instead the lidar is mounted on the aircraft floor and points through an opening in the aircraft floor (Fig. 6.1). The fixed beam points almost exactly along the aircraft vertical axis towards the ground, the exact determination of the beam pointing-angle is described in Sec. 6.2.II.

Despite the limited functionality of the prototype version compared to the envisaged scanning version, all intended research purposes (nadir and wind profiling) can be investigated to their full extent. On the one hand, nadir observations are provided through the default setting of

Fig. 6.1.: Prototype ADL installation inside the Dornier 128-6 aircraft used for the research flights analyzed in this work. (a) Lidar transceiver mounted on the aircraft floor, pointing along the aircraft vertical axis through an opening. Real-time advanced signal processing (RASP) unit for data acquisition mounted in rack. (b) Operator position. Mounted in rack are: work station to control all components, system control computer (SCC), transceiver control unit (TCU) and chiller for lidar cooling (Chiller).

the prototype system with the aircraft in straight and level flight. On the other hand, wind profiling can be conducted by using the aircraft as a scanner, e.g. by banking the aircraft in a turn. Although the spatial resolution of such wind profiles is coarsened compared to a scanning version, all basic measurement principles and the measurement accuracy can be investigated.
The obtained data differs slightly from that of a scanning system. On the one hand, due to the fixed position of the beam, no scanner movement and scanner position need to be considered, and no scanner clock needs to be synchronized. On the other hand, a number of features available with the intended scanning system cannot be utilized. For example, the envisaged scanner is designed to enable real-time aircraft movement compensation. Thereby, the scanning system would be able to point nadir at all times, independent of the aircraft orientation. As the prototype version can deviate from nadir, the effects introduced thereby need to be considered, e.g. for variance measurements of the vertical wind in the BL (Sec. 6.2.VI). The differences, both in data processing and measurement quality, between a fixed beam versus scanning system are discussed in the following sections wherever necessary.

During the four prototype test flights the Do128 operated out of Brunswick airport, consequently the measurement flights covered the surrounding regions. All four flights were successful and are analyzed in the following sections. An overview of the flight tracks is provided in Figs. C.15 - C.18, an overview of the ADL measured data in Figs. C.19 - C.22. For overall measurement quality analysis, the flights are divided into individual legs to distinguish between nadir and turn segments. Further, level flight is distinguished from ascending or descending flight both for nadir and turns. The full set of flight legs analyzed for measurement quality evaluation in the next section is given in Tables B.2 - B.5. All flights contained in-situ atmospheric soundings conducted by the aircraft, which are used for accuracy evaluation of the ADL retrieved wind profiles. The first three of the four flights also contained pitch-roll-yaw (PRY) maneuvers, which can be used to gain detailed insight into the INS motion correction accuracy. Some additional experimental flight maneuvers, which do not fall into any of the previous categories, were also conducted and are not analyzed here. These are labeled as category 'other', and include e.g. short transfers without specific measurement strategy or a Lagrangian cloud-updraft tracking attempt at the end of flight 4.

An initial test flight was conducted on 10 July 2019 in order to demonstrate the basic functionality of all components in flight. Beside this, the focus of the flight was on wind profiling accuracy evaluation, but also incorporated two extended nadir legs over flat terrain. On the first nadir leg the aircraft was flying above an almost cloud-free BL section. Shallow accretions of cumulus humilis were present elsewhere and a thin cirrus was present above (analyzed in Sec. 6.5, see Fig. C.33 for an impression of the cloud field at the beginning of the flight). During the course of the flight, the cirrus was further advected from northwestern Germany over the measurement region and thickened. The second nadir leg was conducted inside the BL along a similar track, again with some shallow cumulus humilis present above the aircraft and a thickened cirrus layer at high altitudes. The BL height during this flight was around 1100 m above mean sea level (MSL, which corresponds to 1000 m above ground level, AGL, all heights are given in MSL in the following unless specified otherwise), with limited further growth, potentially due to the advancing cirrus.

The second and third flight were both conducted on 18 July 2019. During the day, a small upper-tropospheric ridge passed over Germany and a subsequent weak cold front (associated with the following upper-tropospheric trough) slowly approached Germany from west. Preceding the cold front, a convergence line developed over western Germany and advanced eastward during the day, influencing the conditions in the investigation area. Linked with the progressing convergence line, the horizontal wind field in the measurement region was complex and evolving throughout the day, a snapshot from an NWP model is provided in Fig. C.23. First, a morn-

ing flight (flight 2) was conducted across the moderately complex terrain of the Harz mountain region to capture spatial differences in the BL development during the onset of convection in the morning (take-off 3 hours after sunrise). During the first part of the flight, which consists of long nadir transects at high altitude, interrupted by wind profiling turns every 10 minutes, cloud-free conditions and a shallow BL prevailed. Towards the end, with a higher BL top and the development of convective clouds (cumulus humilis) above the Harz, some legs inside the BL (below cloud base) were also conducted.

Second, a similar flight pattern was repeated on the afternoon of the same day (flight 3). During this flight laser stability issues were encountered after take-off, which led to a data loss of 30 minutes flight time ($\approx 20\%$ of the flight time). The instability of the laser after take-off was most likely due to a too high laser temperature on the ground. The aircraft was parked on the tarmac for two hours before take-off, which allowed for substantial heat build-up in the cabin due to insolation. Hence, pre-cooling of the laser unit while on the ground is recommended for future deployments. For flight 3, both synoptic forcing as well as temporal evolution led to a noticeably altered BL compared to flight 2. Extended cumulus humilis were present, which developed to cumulus mediocris above the Harz. The observed cloud bases were in the region 1500 – 2000 m and the cloud tops in the region 2500 – 3000 m. Following this (and due to the VFR limitation of staying below 3000 m), only one extended flight leg could be conducted above the BL. The rest of the flight legs was conducted with the aircraft flying in close proximity to the cloud base and the lidar measuring below. Thereby, the vertical structure of below-cloud updrafts can be captured and differences in the wind field over moderately complex terrain and the surrounding mostly flat terrain can be analyzed. The two flights on 18 July 2019 are analyzed in Sec. 6.6.

A fourth flight was conducted on 19 July before the passage of a weak, partially occluded cold front (which continued its approach to Germany from the previous day and was associated with an advancing trough structure). During this flight the observed wind speeds were higher than on previous flights. The flight was conducted over flat terrain, where extended fields of cumulus mediocris were present, developing into cumulus congestus in a few instances. The cloud base was around 1200 m, whereas cloud tops were around 3000 m but also reaching higher in some places. Two extended nadir transects above the clouds were conducted, again with intermediate wind profiling. The rest of the flight was flown below cloud base and also contained an in-situ self-validation of one nadir transect (utilized in Sec. 6.2.VII). Due to the higher wind speeds, the difference in turbulence structure between along- versus across-wind sampling can be analyzed (Sec. 6.7).

Overall, the diverse meteorological conditions encountered on the four flights are favorable for assessment of the ADL measurement accuracy and potential for a range of measurement and

atmospheric conditions, as well as a number of measurement scenarios, which are presented in the following.

6.2. Measurement processing and quality evaluation

The real-world ADL measurement platform is much more complex than the idealized platform in the ADLS. Consequently, a number of procedures not present in the ADLS have to be implemented in order to achieve high measurement quality. Corresponding with these synchronization, calibration and correction procedures, a number of additional sources of error and associated uncertainties have to be evaluated to control the performance of the procedures. In order to achieve real data synchronization, calibration, correction as well as accuracy assessment I follow the approach of existing studies, which are combined, refined and extended as a part of this work. Before continuing, if not done already, the reader is advised to read Sec. 3.5 of Chap. 3 for an overview of the concepts and terminology used in the following.
In this section, it is evaluated to what degree the mentioned sources of error discussed in Sec. 3.5 can be quantified through conservative error assessments as well as uncertainty estimation methods in real-world measurements. The influence of each error source on the first- and second-order moment is quantified. Thereby, the results also allow a discussion of the validity of the assumption to neglect the effects applicable to real-world ADL measurements in the ADLS simulations in the first place.

6.2.I. Time synchronization

The data acquisition of the Do128 aircraft and the WTX lidar are physically separate systems and do not share a common time server to synchronize the internal system clocks during flight. Thereby, time synchronization between the internal aircraft and lidar clocks has to be performed as a post-processing step, which yields the time difference Δt between aircraft and lidar data acquisition. The time synchronization needs to deal with the problem that the measurement frequency of the aircraft INS data is 100 Hz, whereas the effective measurement frequency of the lidar is 10 Hz. Further, the time synchronization should provide high temporal accuracy. It is shown that despite the lower lidar measurement frequency, time synchronization can be achieved with an accuracy better than 0.1 s. Finally, the time synchronization has to enable control of the temporal variation of the aircraft and lidar clock offset. It is shown that the time synchronization offset is stable over the duration of one flight, whereas for different flights the offset may vary.

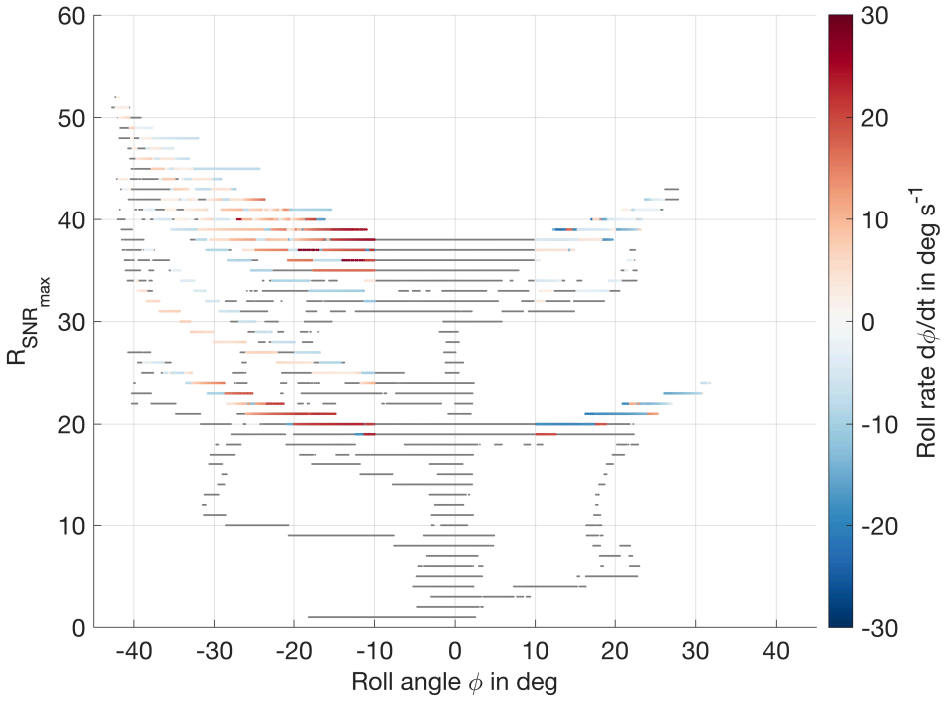

Fig. 6.2.: Illustration of lidar sensing distance as a function of aircraft roll angle. The number of range gate with maximum SNR is displayed as a function of aircraft roll angle ϕ for the morning flight on 18 July 2019. The roll rate is color coded, gray values show data which do not fulfill the time synchronization criteria.

Time synchronization procedure

In this work, time synchronization is achieved based on a newly developed procedure. The procedure exploits the variation in the maximum lidar sensing distance (e.g. distance to the ground) during aircraft maneuvers. For example, an increase in aircraft roll angle causes an increase in slant path length of the lidar beam in the atmosphere, linking the independent measurements of the two systems. The time synchronization procedure is demonstrated for the morning flight of 18 July 2019, but applied to all other flights in a similar manner.

Before time synchronization, a rough value of the time offset with accuracy $< 2\,\mathrm{s}$ is obtained by visual analysis of aircraft and lidar data. Visual determination is easily possible using the measured radial velocities, for example during turns or PRY maneuvers (Appendix A.6). At the beginning of the time synchronization procedure, the maximum SNR value along the beam is determined for each individual lidar measurement. The number of the range gate in which the maximum SNR value occurs is stored as $R_{SNR_{max}}$. This data serves as a proxy for ground detection by the lidar (it is independent from the aircraft data). The data is quality controlled by processing only values with SNR $> 5\,\mathrm{dB}$, further, in order to exclude clouds and other erroneous measurements, the number of the detected ground range gate needs to be within ±2 of

the ground range gate detected using a digital elevation model (DEM, Appendix A.5). Then, the 10 Hz $R_{SNR_{max}}$ data is interpolated to 100 Hz by means of nearest neighbor interpolation. Although the interpolated data is physically not meaningful (and still digitized) it is useful in order to enable calculations using both the aircraft and lidar data set. The aircraft data is also filtered for defined roll movements (which exhibit large magnitude and thereby show the stronger signal compared to pitch and raw movements), only data with roll angle $\phi > 10°$, roll-rate $|d\phi/dt| > 2\,°\,s^{-1}$ and 60 s averaged vertical aircraft velocity $|dH_{60s}/dt| < 1\,m\,s^{-1}$ are considered. These settings have to be adapted should the procedure be applied to another type of aircraft.

Figure 6.2 shows $R_{SNR_{max}}$ for every measurement as a function of the aircraft roll angle ϕ. As expected, for an increase in aircraft roll angle, $R_{SNR_{max}}$ also increases, yielding a wine-glass shape (at different levels, as turns are conducted at different altitudes). Assuming a fixed beam installation pointing nadir in ARD (a valid approximation, as is shown in Sec. 6.2.II), the length of the lidar beam slant path in the atmosphere due to roll is proportional to $1/b_z^E$ ($1/cos(|\phi|)$ if only the roll angle is considered), which is used to linearize the relationship. Thus, a linear correlation coefficient can be calculated for each detected roll movement between the aircraft attitude angles and the lidar measured ground range gate number $R_{SNR_{max}}$.

In order to obtain robust results, roll movements are detected as full oscillations, meaning both the increase in absolute roll angle (e.g. beginning of turn) and decrease in roll angle (e.g. end of turn) are used for calculation of the correlation coefficient. Figure C.24 provides a display of the used data as a function of time, including the detected begin and end time of each turn segment. Roll maneuvers during each PRY maneuver (multiple oscillations) are combined and used together. During the morning flight on 18 July 2019 PRY maneuvers were conducted at the beginning of the flight (06:33:20-06:35:20 UTC), the middle of the flight (07:41:55-07:43:50 UTC) and twice at the end of the flight (08:37:00-08:38:30 UTC, 08:38:50-08:40:15 UTC) in order to gain insight into the stability of the time offset between aircraft and lidar system.

The high accuracy time offset $\Delta t_{Do128-WTX,raw}$ is determined through a correlation analysis in which the aircraft data is lagged against the lidar data. The procedure is conducted for all detected roll maneuvers on the interval $-1\,s < \Delta t_{Do128-WTX} < 1\,s$ iteratively, until the best agreement is reached (e.g. maximum correlation at lag zero). The result of the lagged correlation analysis, using an optimum $\Delta t_{Do128-WTX,raw} = 18.12\,s$, is displayed in Fig. C.25. Most curves show a maximum on the investigated interval, especially the correlation analysis obtained from the four PRY segments show a distinct maximum. In the following, the vertical distance between the maxima and minima of the correlation curve on the investigated interval, the so called correlation peak prominence, is used to identify the reliability of the time synchro-

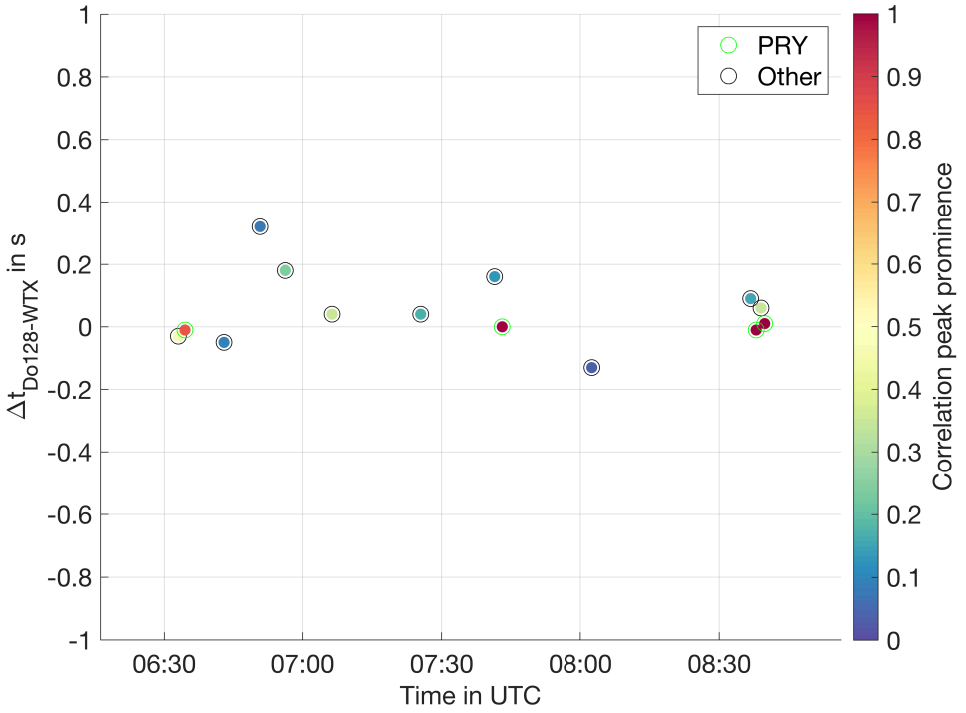

Fig. 6.3.: Time difference $\Delta t_{Do128-WTX}$ at which the correlation curves from Fig. C.25 are maximal as a function of flight time for the morning flight on 18 July 2019. The raw time difference applied is $\Delta t_{Do128-WTX,raw} = 18.12\,\text{s}$. The correlation peak prominence on the investigated interval is color coded. Values obtained from PRY maneuvers are circled in green, values obtained from other maneuvers are circled in black.

nization achieved. Confirming the suitability of the time synchronization procedure, it should be noted that the localization of the maxima has higher time resolution than the raw lidar data which exhibits only 0.1 s resolution. Th higher time resolution is enables by the interpolation of the lidar data to higher measurement frequency, while maintaining the stability and accuracy of the solution due to the exploitation of low frequency roll movements. Correlation curves which do not show a maximum on the investigated interval are linked to either a non-sufficient detection of full turn-segments, deficiencies of the aircraft INS data during maneuvers or an imperfect ground range gate identification. Due to the well-localized maxima, time synchronization from PRY maneuvers is more reliable and preferable.

Time synchronization results

The time offset stability over the course of the flight can be extracted from Fig. 6.3. Due to the less prominent maxima, analysis of turns and other maneuvers yields higher scatter and is less reliable for time synchronization (although far from being unusable). The PRY maneuvers yield stable results for the time difference $\Delta t_{Do128-WTX,raw}$, with prominent maxima localization.

The difference between the time synchronization offsets determined from the individual PRY maneuvers is $< 0.03\,\mathrm{s}$, much smaller than the lidar measurement resolution. Based on this analysis, it is clear that the time synchronization offset between aircraft and lidar system is constant over the duration of one flight.

The procedure illustrated above is repeated for all flights, yielding the time synchronization offsets displayed in Table 6.1. Based on the time synchronization accuracy achieved during PRY maneuvers, the remaining time synchronization uncertainty is conservatively estimated to be $< 0.1\,\mathrm{s}$.

Tab. 6.1.: Time synchronization offsets between aircraft and lidar for each flight.

Flight number	$\Delta t_{Do128-WTX,raw}$
1 - 10 July 2019	18.37 s
2 - 18 July 2019, morning	18.12 s
3 - 18 July 2019, afternoon	18.10 s
4 - 19 July 2019	18.17 s

Influence of time synchronization on mean and variance of motion-corrected radial velocity

During stable flight, aircraft movements cancel out over short intervals of time, therefore the influence of an imperfect time synchronization cancels out on average, as times with an overestimation or underestimation occur with equal likelihood. Consequently, an imperfect time synchronization does not lead to a systematic uncertainty in the estimated first order moment. The random uncertainty is estimated based on the average aircraft velocity change during a 0.1 s interval, its value below $1 \times 10^{-4}\,\mathrm{m\,s^{-1}}$ does not require further consideration.

For the lidar configuration presented here, the maximum additional variance (systematic influence on the second order moment) introduced due to imperfect time synchronization is the aircraft velocity variance during a 0.1 s interval. As the lidar points almost straight down, the variance in aircraft velocity along the aircraft z-axis is relevant, it is determined to be $\overline{\boldsymbol{v}_{a\ z,0.1s}^2} < 0.005\,\mathrm{m^2\,s^{-2}}$ for all measurement flights. As a result, the uncertainty in the lidar measured variance due to time synchronization is thereby conservatively estimated as $\overline{\beta_t^2} < 0.005\,\mathrm{m^2\,s^{-2}}$. A leg specific maximum time synchronization uncertainty can also be calculated for analysis based on the same assumptions and theory.

Scanning system changes

For a scanning ADL time synchronization between lidar and scanner also becomes important. Due to the fast scan rates accomplished by the scanner, as well as the much larger horizontal aircraft speed being projected into the measurement for horizontal scans, the possible error due to insufficient time synchronization is much larger. However, for scanning systems, an even

more accurate synchronization is achievable through exploiting the projection of the horizontal aircraft onto the beam itself, based on a slightly modified version of the above correlation technique. Should a (desirable) common time server be available for all instruments the procedure discussed here can be used to evaluate the reliability of the time distribution to the individual clients.

6.2.II. Beam alignment

The beam pointing-angle needs to be determined accurately for reliable measurements. An erroneous beam pointing-angle will result in wind measurement offsets depending on the angle between the aircraft heading and the wind direction. The beam pointing-angle cannot be determined with sufficient accuracy (<0.5°) through examining the geometric installation of the lidar in the aircraft. Consequently, a beam pointing calibration is conducted based on the procedure developed by Haimov and Rodi (2013, for ADR) as a part of this work.

Beam pointing-angle calibration procedure

The beam pointing calibration procedure developed by Haimov and Rodi (2013) exploits the fact that the theoretical ground return velocity is $\boldsymbol{v}_p = 0\,\mathrm{m\,s^{-1}}$, as the ground is a stationary target. Therefore, the measured ground return velocity should come to zero after correction of the measured data for the aircraft motion contribution using the aircraft INS data. Systematic remaining ground return velocities are associated with a beam pointing-angle offset, thereby the ground return velocities can be used to calibrate the beam pointing-angle. Short-term errors in the ground return velocities cancel out when using long flight times and multiple flights segments from different flights. Therefore, the procedure can be applied despite the presence of short-term errors in the ground return velocities due to the other measurement errors (discussed in Sec. 6.2.III and Sec. 6.2.IV, as well as errors associated with identifying the ground return, see Appendix A.5). For the calibration procedure, the equation system (Eq. 3.12) is reformulated for N measurements of the ground return velocity (assuming the theoretical $\boldsymbol{v}_p = 0\,\mathrm{m\,s^{-1}}$), with the beam pointing direction $\boldsymbol{b} = [b_x^A, b_y^A, b_z^A]$ as calibration constants, and $\boldsymbol{v}_L{}^A$ according to Eq. 3.10, giving:

$$v_{Di} = b_x^A v_{Lx,i}^A + b_y^A v_{Ly,i}^A + b_z^A v_{Lz,i}^A, \quad i = 1, 2, 3, ..., N \tag{6.1}$$

and the additional constraint

$$\sqrt{(b^A_{x,i})^2 + (b^A_{y,i})^2 + (b^A_{z,i})^2} = 1. \tag{6.2}$$

The second condition is necessary and useful, as the determination of the beam pointing direction in y-direction is prone to larger uncertainty (Haimov and Rodi, 2013). The larger uncertainty results from the smaller beam direction component in y-direction as well as smaller projection of aircraft velocities in this direction. Typical ARD aircraft velocity components in y-direction do not exceed $5\,\mathrm{m\,s^{-1}}$, due to the small crosswind component compared to the aircraft velocity vector and small side-slip angles during normal measurement flights (for a directional cosine of the y-direction by less than 0.5° from 90° the $\boldsymbol{v}_a$ y-component is less than $0.1\,\mathrm{m\,s^{-1}}$ and therefore difficult to determine accurately).
The calibration is performed in ARD, as the aircraft moment-arm contribution is simple in this coordinate system compared to ENU. The beam pointing calibration conducted utilizes a ground return identification based on DEM ground elevation data (Appendix A.5). For solution of the equation system the MatLab least-square solver 'fmincon' is employed in standard setup, other solvers were also tested and give similar results. All segments from all flights are analyzed for beam pointing-angle calibration, as the lidar installation was not changed between flights (see Tab. B.2 - B.5 for an overview of all flight segments analyzed).

Beam pointing-angle calibration results

The results obtained from the beam pointing-angle calibration are displayed in Fig. 6.4. In line with the results obtained by Haimov and Rodi (2013), due to the reasons outlined above, the calibration of the b_x, b_z beam components yields a sharper distribution with less outliers, compared to the b_y direction. b_x shows by far the sharpest distribution, due to the large magnitude of the aircraft velocity vector in this direction. The b_z component shows more scatter, as the aircraft velocity component in the vertical direction is more difficult to determine (Lenschow and Sun, 2007). Due to this difficulty, beam calibration results from ascending and descending legs also show degraded accuracy. Contrary to the results obtained by Haimov and Rodi (2013), the calibration results obtained from turns are less accurate compared to those from nadir legs. As discussed by them, their greater nadir uncertainty is because their nadir legs were flown in along-wind direction frequently, which is not the case in this work. However, the beam calibration results obtained from turns should yield a more accurate determination of the beam pointing angle than observed here, which points to problems with the INS used. This conclusion is further manifested by the observation that the b_y direction shows an almost linear

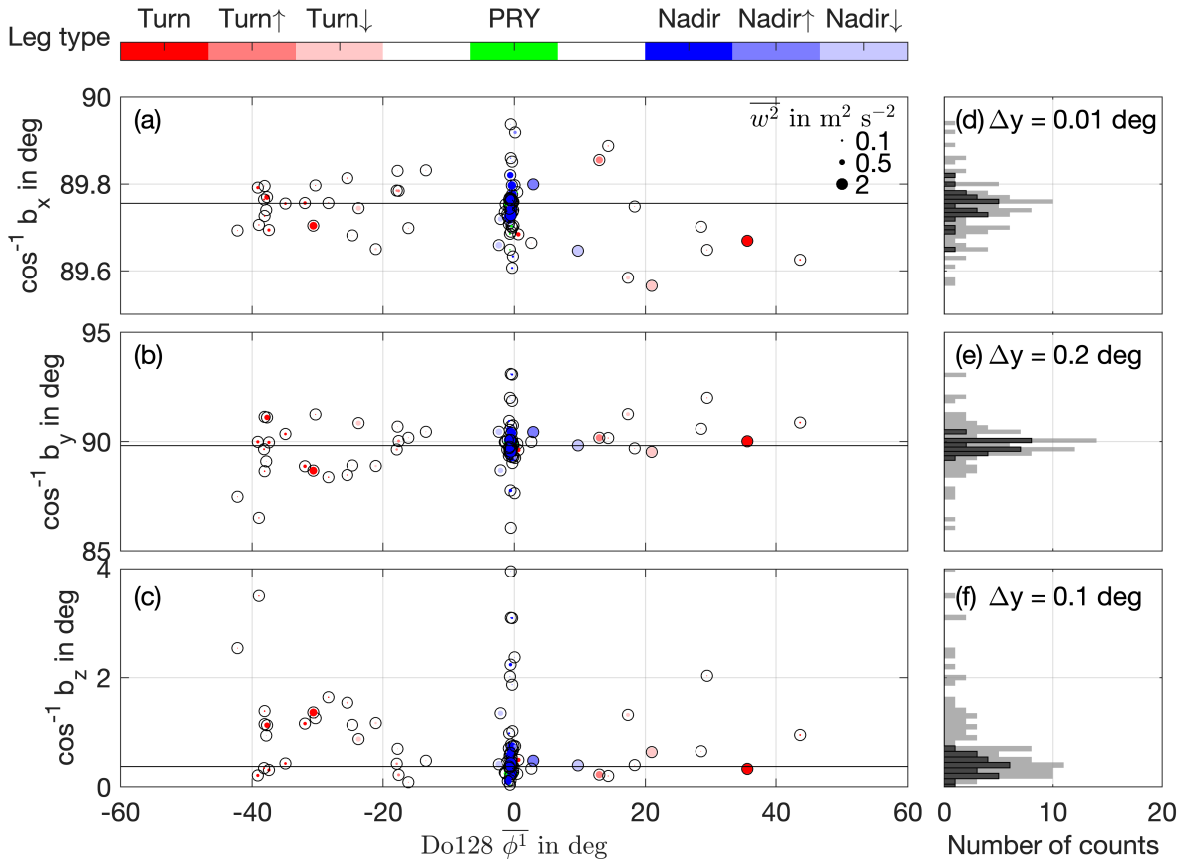

Fig. 6.4.: Results of beam pointing-angle calibration as a function of average roll angle. (a), (b), (c) beam pointing directional cosines obtained from calibration for b_x, b_y, b_z. The leg type is color coded and the average atmospheric variance $\overline{w^2}$ measured by the aircraft is displayed through the circle area. The horizontal black lines represent the median values used as calibrated beam pointing-angles. Please note the differences in y-axis scaling. (d), (e), (f) beam pointing directional cosine histograms for b_x, b_y, b_z. Grey histograms show the original distributions for all legs, black histograms show the quality controlled distributions according to the specifications given in the text. Please note the different bin-widths used for calculating the histograms.

change as a function of roll angle, with opposite behavior for right vs. left turns. A systematic temporal evolution of the beam pointing angle during the course of a flight is not observed, hinting against the Schuler-oscillation as the main driver of the observed INS deficiency. A faulty aircraft moment-arm contribution correction is also ruled out, as the raw ground return velocities show no systematic error during turns (which they do if the moment-arm contribution correction is turned off for test purposes). In summary, the observations made here point to an inferior quality of the INS onboard the Do128 compared to the INS available in Haimov and Rodi (2013), especially during turn and non-constant level maneuvers (see also Sec. 6.2.IV). Nevertheless, the Do128 INS quality is sufficient to enable reliable beam pointing-angle calibration, as is shown by the well-defined histograms, which can be further quality controlled. Due

Tab. 6.2.: Beam pointing-angles and associated uncertainties obtained from quality controlled calibration.

	median in deg	$\overline{\varepsilon^1}$ in deg
$\cos^{-1} b_x$	89.756	0.036
$\cos^{-1} b_y$	89.815	0.301
$\cos^{-1} b_z$	0.373	0.170

to the problem during turns, only the results from PRY maneuvers and straight and level nadir legs are used for calibration (for a description of PRY maneuvers see Appendix A.6). For nadir legs, sufficient exploration of the b_x, b_y and b_z directions needs to be assured to avoid an ill-posed problem with larger condition number. For this reason, only calibration values obtained from flight in a disturbed atmosphere are used and a threshold $\overline{w^2} > 0.5\,\mathrm{m^2 s^{-2}}$ (as measured by the aircraft) is applied for nadir legs.

The average calibrated beam pointing-angle is calculated as the median of the quality controlled distribution. Obtained values of the beam directional cosines and their associated uncertainties are given in Table 6.2. Combining the uncertainties in the least-favorable way results in a maximum beam pointing-angle uncertainty of $\delta_b = 0.303°$, largely driven by the cross-beam directional uncertainty (and smaller than the algebraic addition of the uncertainties because the uncertainties in y, z direction are strongly co-dependent). Calculating the beam pointing-angle separately for straight and level nadir vs. PRY legs yields very similar solution, giving further confidence in the obtained results. The calibrated beam pointing-angles given in Table 6.2 correspond to beam vector components of (the deviation of the vector magnitude from 1 is due to round-off error of the shown values):

$$\begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = \begin{bmatrix} 0.0043 \\ 0.0032 \\ 0.9999 \end{bmatrix}. \tag{6.3}$$

Influence of beam pointing-angle on mean and variance of motion-corrected radial velocity

The erroneous component in the motion-corrected radial velocity due to the remaining beam pointing-angle uncertainty can be estimated in ARD based on Damiani and Haimov (2006):

$$\overline{\beta_b^1} = (\boldsymbol{b}^A - \boldsymbol{b}_\varepsilon^A) \cdot \boldsymbol{v}_a^A. \tag{6.4}$$

$\boldsymbol{b}_\varepsilon$ is a conservative estimate of the beam pointing-angle uncertainty, obtained through adding or subtracting one standard deviation of the beam pointing-angle uncertainty in the least-favorable way (assuming a constant beam pointing-angle offset). Compared to Damiani and Haimov (2006), the above formulation neglects the shear term of the horizontal velocity in the cross-beam direction. For the single Doppler retrieval presented here, a potential cross-beam displacement of the beam under the presence of shear simply measures another region of the wind field but does not influence the retrieval accuracy, as no dual-Doppler analysis is conducted. Further, the uncertainty due to an erroneous aircraft motion correction is also excluded, as this effect is analyzed separately in Sec. 6.2.IV.

As discussed above, an error in the beam pointing-angle results in an error in the motion-corrected radial velocity depending on the angle between aircraft heading and wind direction. If the beam pointing angle error can be assumed to be constant, the erroneously introduced velocity in ARD depends only on the aircraft velocity vector encountered during the flight leg. As the Do128 cabin is unpressurized, aircraft frame distortion due to changing cabin pressure is not a problem compared to the Beechcraft Super King Air 200T used by Haimov and Rodi (2013), therefore the beam pointing angle is unaffected by flight level changes. Further, distortion of the aircraft frame due to external forces (turbulence, pilot maneuvers) plays a minor role for the beam pointing-angle, based on the results by Haimov and Rodi (2013) and the observed minor increase in ground return velocity variance in turbulent flight conditions (Fig. 6.11). In conclusion, the beam pointing-angle error can be assumed to be constant for all flights.

Using the maximum values of $\boldsymbol{v}_a^A = [92.29, 20.02, 16.17]\,\mathrm{m\,s^{-1}}$ encountered during any flight leg and the least-favorable beam pointing uncertainty, the maximum random velocity introduced is estimated as $\overline{\varepsilon_b^1} < 0.17\,\mathrm{m\,s^{-1}}$ for the mean. This maximum velocity can appear as a constant offset in the motion corrected radial velocity for individual legs. However, the effect on the retrieved wind (in ENU) varies depending on the angle between aircraft heading and wind direction. Hence, it may vary between subsequent legs, mandating its description as a random uncertainty for the overall accuracy assessment of the mean. Additionally, the above method allows for calculation of a leg-based uncertainty, which is useful as the uncertainty is reduced for straight and level nadir legs with smaller aircraft velocities in the y- and z-direction. The potential systematic offset due to remaining beam-pointing angle uncertainty is estimated based on the maximum measured systematic offset of the (non-quality controlled) ground return velocity during straight and level maneuvers over any flight, which is given as $\overline{\beta_b^1} < 0.05\,\mathrm{m\,s^{-1}}$ (Appendix A.5). Based on the measured ground return velocities, the values given here are shown to be strong overestimations in Sec. 6.2.IV.

The maximum effect on the measured variance during straight and level nadir legs is calculated as $\overline{\beta_b^2} < 0.033\,\mathrm{m^2\,s^{-2}}$, which is based on the maximum variance in the aircraft velocity vector encountered during any straight and level nadir leg ($\overline{V_a^2} < [12.01, 4.92, 2.12]\ \mathrm{m^2\,s^{-2}}$) and the most unfavorable beam pointing uncertainty. For the variance measured during turns the contribution is estimated to be $\overline{\beta_b^2} < 0.37\,\mathrm{m^2\,s^{-2}}$, again calculated from the maximum variance in the aircraft velocity vector encountered during any turn ($\overline{V_a^2} < [90.17, 53.34, 25.79]\ \mathrm{m^2\,s^{-2}}$). As for the mean, the conservative values given here are shown to be strong overestimations based on the measured ground return velocities in Sec. 6.2.IV.

Scanning system changes

An accurate beam pointing-angle calibration is even more important for a scanning ADL system compared to the fixed beam system investigated here, due to the approximately one order of magnitude larger aircraft velocity components introduced into the radial velocity measurements. However, due to the larger aircraft velocity components, the calibration procedure is also expected to yield an order of magnitude more accurate results, yielding an overall similar measurement accuracy.

6.2.III. Lidar signal quality control

Quality control of the lidar data involves two closely linked procedures. A quality control mask has to be applied in order to exclude unreliable lidar measurements, which can be caused by too low SNR or hard targets such as clouds. In this work, a two-pass SNR-based quality control mask is applied, with additional checks for consistency and plausibility (similar techniques have been used by Gamache, 2005, and Weissmann et al., 2005b). The ground is identified using both the signal to noise ratio and a DEM model. Clouds are detected through their hard target characteristics at altitudes different from the ground. In this section, first the functionality of the quality control mask is explained. The specific SNR thresholds used by the quality control mask are determined based on an uncorrelated noise estimation, which is explained thereafter. The two procedures are intertwined and developed iteratively in practice, as on the one hand the quality control mask needs SNR-thresholds determined from uncorrelated noise estimation. On the other hand, the uncorrelated noise estimation works most reliably with quality controlled data and provides a cross-check on the quality obtained.

Quality control mask

The quality control procedure is demonstrated for the morning flight on 18 July 2019 using data from the first range gate (Fig. 6.5). The flight was conducted at high altitude (up to 3000 m), thereby presenting challenging conditions due to the low SNR in the free troposphere. The first range gate is both highest up as well as the most difficult to process, as measurement artifacts from the lidar blind zone are present sometimes (e.g. Fig. 6.5 shows a lot of measurements scattered around $0\,\mathrm{m\,s^{-1}}$).

The quality control procedure is implemented as follows: First, high quality measurement time series are created for all range gates using a strict quality control criteria of $SNR > -5\,\mathrm{dB}$ (this threshold is motivated based on the uncorrelated noise results). A few remaining unreliable measurements remain even at high SNR, these are excluded by applying a jump filter eliminating subsequent measurements with a velocity difference $|\Delta V_D| > 5\,\mathrm{m\,s^{-1}}$. After this procedure, a time-dependent valid measurement range is determined based on the high quality measurements (if any are available, otherwise the data is not used). To this end, for every time the range is calculated as the minimum and maximum radial velocity obtained from *all* range gates on a $10-$s interval (the $10-$s moving minimum and maximum envelope). This procedure yields a time varying valid measurement corridor, shown by the blue lines in Fig. 6.5. Subsequently, all measurements present in this corridor and fulfilling a relaxed SNR quality control criteria $> -15\,\mathrm{dB}$ are considered. The final, quality controlled time series is obtained through application of a spike filter eliminating spikes with a peak $|\Delta V_D| > 2\,\mathrm{m\,s^{-1}}$ compared to the preceding and subsequent neighboring measurements (the red points in Fig. 6.5). If the neighboring values are missing, an artificial radial velocity level is constructed as the running median on a $10-$s interval, similar to Kiemle et al. (2007) but only for the purpose of spike detection. The spike filter removes measurements which fall into the valid measurement corridor but are random in nature. Visual analysis of the filtered time series for all flights and range gates ensures that the applied spike filter does not remove any data during measurements in the BL. Figure 6.5 shows the added value of the two-level quality control procedure. A great number of additional measurements are available in the valid measurement corridor in regions of low SNR (e.g 07:15 UTC - 07:45 UTC). The spike filter removes points which are unreliable or wrongly included due to a too large corridor interval (which also reveals that some of the faulty $0\,\mathrm{m\,s^{-1}}$ values show elevated $SNR > -5\,\mathrm{dB}$, e.g 06:45 UTC - 06:50 UTC). The additionally obtained measurements are less important for nadir analysis (as they are located in the free troposphere, where turbulence levels are low), but more important for wind profiling purposes, where additional retrieved wind profile points become available in the free troposphere.

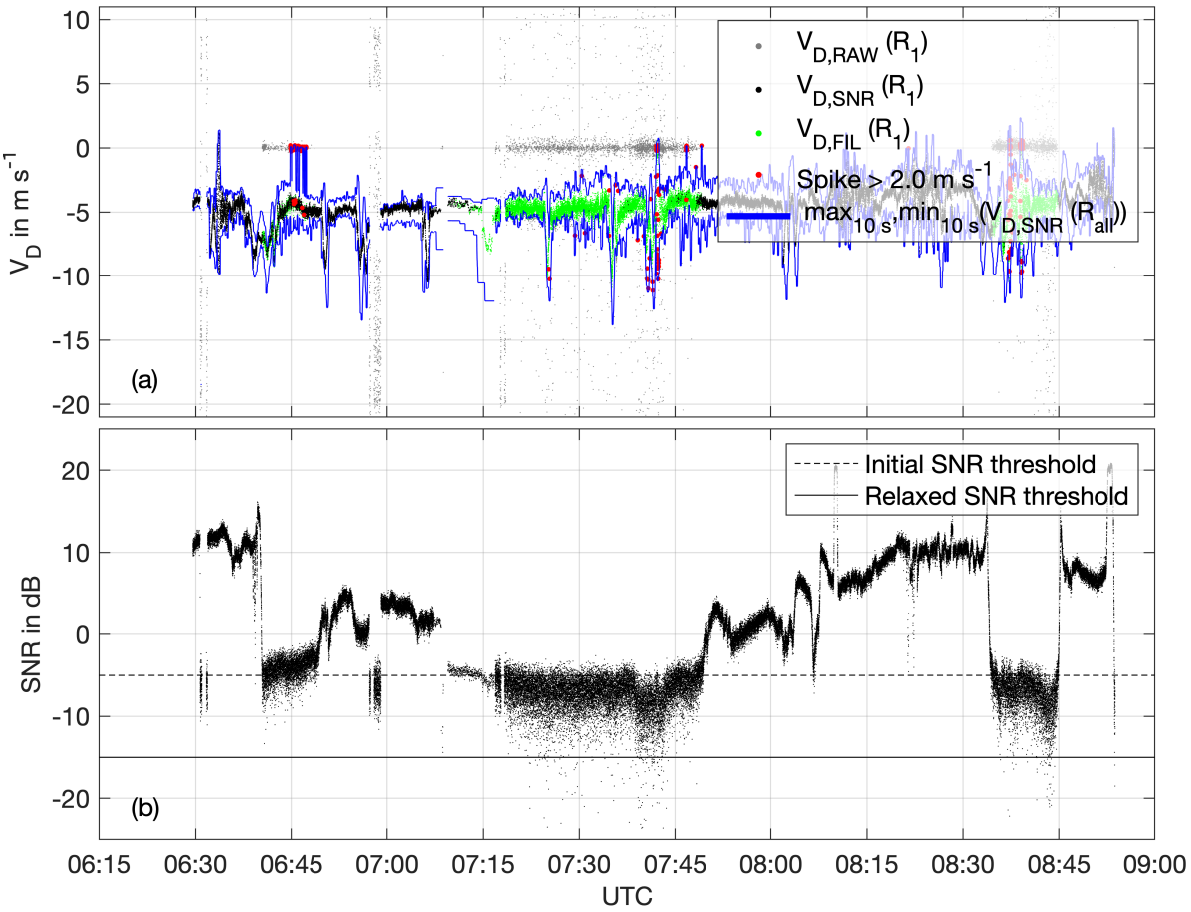

Fig. 6.5.: Illustration of quality control procedure for the first range gate during the morning flight on 18 July 2019. (a) shows the raw Doppler velocities (gray), the strict SNR quality controlled Doppler velocities (black, −5 dB threshold) and the additional Doppler velocities obtained by the second pass (green). Additional velocities are obtained by thresholding in a 10 −s moving minimum-maximum limited measurement corridor (blue lines), based on the strict SNR quality controlled Doppler velocities from all range gates. Outliers detected as spikes are shown in red. (b) shows the associated SNR values.

After the basic quality control routine additional quality flags are set for the data. The range gate intersecting with the ground is determined using a DEM model for ground elevation (Sec. A.5). The measurements of the two range gates above ground intersection can be compromised by chirp effects and are flagged for unreliability (as the ground return velocity is an important control parameter in the following sections it is only flagged for meteorological analysis but not excluded during data processing in general).

Clouds are determined by evaluating the cross-check between the DEM determined ground range gate and the range gate exhibiting maximum SNR and the maximum backscatter gradient, individually for all beams. If the numbers of the range gate with maximum SNR and maximum backscatter gradient agree within ±3 a hard target detection is confirmed. If the confirmed hard target detection differs by more than 3 range gates from the ground range gate determined

using the DEM, the hard target detection is associated with a cloud (fog was not present during any of the measurement flights, making this coarse resolution valid). As for the ground, the measurements of the two range gates above the hard target can be compromised by chirp effects and are flagged for unreliability.

Radial velocity estimation errors

Lidar measurements can suffer from both systematic and random errors in the estimated radial velocities. Systematic errors can be caused by hardware issues (e.g. frequency drift of the laser, non-linear amplifiers, digitization errors). Systematic errors have been shown to be small for the type of system used in this work. Frehlich et al. (1994) shows that the systematic uncertainty of a single pulse estimate is $< 0.04\,\mathrm{m\,s^{-1}}$, which becomes $< 0.004\,\mathrm{m\,s^{-1}}$ for a PRF of 750 Hz and averaging of 75 pulses based on Frehlich (1997), it is thereby negligible.
Therefore, random fluctuations of the measured radial velocity remain as an important source of error when measuring atmospheric variance (they are negligible for measurements of the mean as they average out over short time intervals). Important drivers of random fluctuation in the radial velocity are internal laser noise, e.g. due to the photon-limited detector noise, as well as the limited number of scatterers, their random position and turbulence in the measurement volume (Frehlich, 1997, speckle effect, the lidar signal has random amplitude and phase from shot to shot). As the atmospheric decorrelation time for the illuminated aerosol is $< 1\,\mu$s, the random error from consecutive lidar shots is uncorrelated for a PRF of 750 Hz used in this work. Consequently, averaging a number of lidar shots can reduce the uncorrelated noise in the lidar measurements (assuming a frozen wind field over the averaging time). The magnitude of the uncorrelated noise remaining after averaging mainly depends on the SNR. As stated before, for reliable results the uncorrelated noise estimation should be applied after application of the quality control mask, which nevertheless requires SNR-based thresholds already.
The uncorrelated noise can present an important (always positive) source of erroneous variance contribution $\overline{\beta_{rv}^2}$ to the atmospheric variance, especially at low levels of atmospheric variance. Fortunately, a number of methods exist to estimate the systematic uncorrelated noise offset, thereby it is possible to subtract it from the measured variance in order to retain purely the atmospheric variance. In the converse argument, the level of noise remaining can also dictate the number of lidar shots which have to be averaged (and thereby the effective measurement frequency) in order to obtain acceptable levels of uncorrelated noise. The WTX lidar used has a PRF of 750 Hz and standard setups used for the WTX data output rate vary between $1 - 10$ Hz.

Random radial velocity noise - estimation procedure

In this section, the systematic contribution to the measured variance due to random radial velocity noise (uncorrelated noise) is estimated. Three methods exist for estimation of the random radial velocity noise, all three methods are shown to perform comparably by Lenschow et al. (2000) and Frehlich and Cornman (2002). First, the uncorrelated noise can be detected as a jump in the autocovariance function at lag zero (see Sec. 3.4 for the definition of the autocovariance). Two versions exist to quantify the exact magnitude based on this method. The uncorrelated noise can be obtained as the difference between the autocovariance at lag zero minus that at lag one (Lenschow et al., 2000; Stawiarski et al., 2013). This version yields very stable results, as the autocovariance at lag zero is always greater than at lag one, giving a conservative overestimation. However, with this method the uncorrelated noise is overestimated, as the turbulent signal itself can also cause a decay between lag zero and lag one, which is erroneously quantified as uncorrelated noise. For this reason, another common version estimates the level of uncorrelated noise by fitting a linear function through the first lags (usually two to four for 1 Hz measurement data Lenschow et al., 2000, Frehlich, 2001) and extrapolating to lag zero. This version is expected to capture the real magnitude of the uncorrelated noise more reliably under turbulent conditions (smaller overestimation), but only if a concave shape of the autocovariance function at the smallest lags and a well-defined discontinuity at lag zero exist, which might not always be the case (Träumner, 2010). Second, in the spectral method the level of uncorrelated noise is calculated as the constant level of the temporal spectrum at high frequency (Lenschow et al., 2000, see Fig. 6.26 for an illustration). Third, a velocity difference method based on two independent velocity estimates from even and odd numbered pulses can also be employed (Frehlich and Cornman, 2002).
Most studies do not analyze the raw pulse-by-pulse data but only the averaged lidar data output. Consequently the velocity difference method is not widely used and documented, it is therefore also not employed here. The spectral method was found to be more noisy than the other methods by Lenschow et al. (2000), it is therefore also not used here. The analyzed data does not always exhibit concave shaped autocovariance functions at the smallest lags, sometimes producing unrealistic negative values of uncorrelated noise. This problem requires careful uncorrelated noise estimation using both, extrapolation as well as the first lag difference method.

Random radial velocity noise - estimation results

The results of the uncorrelated noise estimation using the first lag method are shown in Fig. 6.6 for the 10 Hz lidar data. All range gates from all legs and all flights are shown, the SNR is

determined as the average SNR over the flight leg at the respective altitude. A reliable uncorrelated noise estimation is only possible for sufficiently long legs, for this reason the leg length is displayed through the circle area and a minimum length of 60 s is required (more specifically, the points display the available time series lengths for the individual range gates).

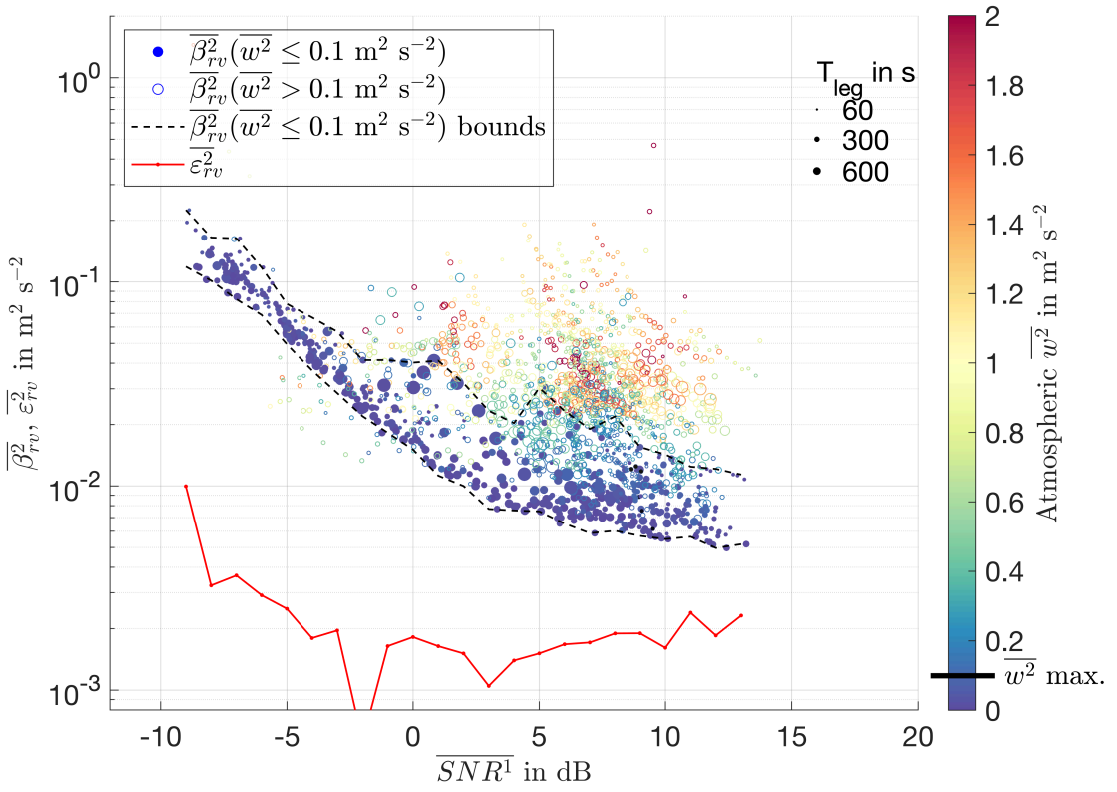

Fig. 6.6.: Uncorrelated noise in lidar signal as a function of average SNR, estimated using the first-lag autocovariance difference method. The atmospheric variance present in the time series used for noise estimation is color coded. Points with an atmospheric variance $\overline{w^2} < 0.1\,\mathrm{m^2\,s^{-2}}$ are filled and their envelope is shown by the dashed black lines. The leg length is displayed through the circle area, a the minimum leg length of 60 s is required for calculation. The red line depicts the random uncertainty of the noise estimation, calculated as the average of the absolute differences in uncorrelated noise between the direct and the linear estimation methods.

Due to the conservative estimation method, increased levels of uncorrelated noise are observed at increased levels of atmospheric variance. This increase vanishes if the linear extrapolation method is used (Fig. 6.7), it is thereby identified as an overestimation due to the first lag estimation method. However, the linear extrapolation method can yield nonphysical negative values of uncorrelated noise, if linear or slightly convex shaped autocovariance functions are present at the first lags. Based on the first lag method, the negative values obtained using the linear extrapolation method can be attributed to convex shaped autocovariance functions for

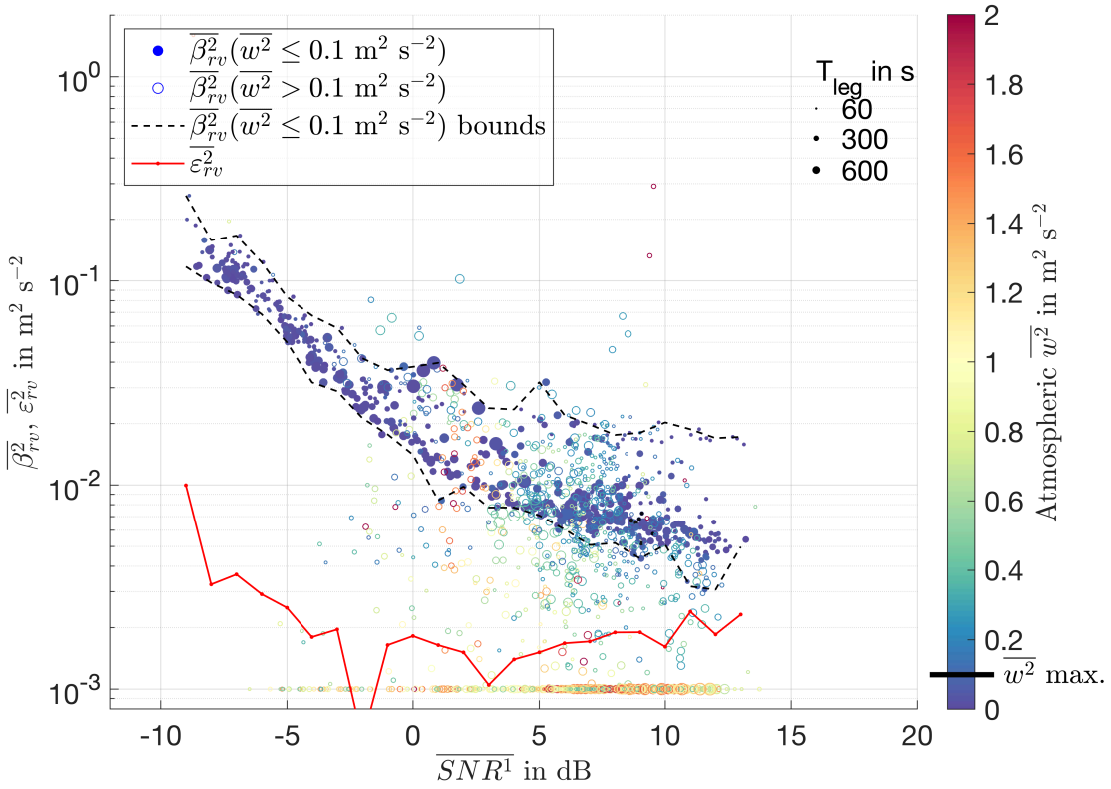

Fig. 6.7.: Uncorrelated noise in lidar signal as a function of average SNR, estimated by extrapolation of the autocovariance function to lag zero using a linear fit through the first four lags. The atmospheric variance present in the time series used for noise estimation is color coded. Points with an atmospheric variance $\overline{w^2} < 0.1\,\mathrm{m^2\,s^{-2}}$ are filled and their envelope is shown by the dashed black lines. Values with uncorrelated noise $\overline{\beta_{rv}^2} < 0.001\,\mathrm{m^2\,s^{-2}}$ are set to $0.001\,\mathrm{m^2\,s^{-2}}$ as these are caused by convex shaped autocovariance functions at the smallest lags. The leg length is displayed through the circle area, a the minimum leg length of 60 s is required for calculation. The red line depicts the random uncertainty, calculated as the average of the absolute differences in uncorrelated noise between the direct and the linear estimation methods.

higher levels of atmospheric variance ($\overline{w^2} > 0.5\,\mathrm{m^2\,s^{-2}}$), in a regime where uncorrelated noise becomes negligible compared to the atmospheric variance.

In conclusion, only values of atmospheric variance $< 0.1\ \mathrm{m^2\,s^{-2}}$ are considered for noise analysis (uncorrelated noise dominated regime). In this regime, the behavior of the uncorrelated noise as a function of SNR becomes clear and produces very similar results for both methods, shown by the dashed black envelope lines. In the uncorrelated noise dominated regime, levels of uncorrelated noise are always smaller than $\overline{\beta_{rv}^2} < 0.3\,\mathrm{m^2\,s^{-2}}$ for all values of SNR, nevertheless, the analysis can be refined further. The uncorrelated noise level remains almost constant at levels $< 0.04\,\mathrm{m^2\,s^{-2}}$ for SNR values larger than 0 dB. For values $-6\,\mathrm{dB} < \mathrm{SNR} < 0\,\mathrm{dB}$ the uncorrelated noise increases more rapidly, while still being far from unusable with levels $< 0.1\,\mathrm{m^2\,s^{-2}}$. The uncorrelated noise level is also well-defined for $\mathrm{SNR} < -6\,\mathrm{dB}$, however, in this

region the levels can exceed those of the atmospheric variance and therefore should be used for variance analysis only with caution. The results clearly show the benefit of the developed quality control procedure, which yields a considerable amount of data with sufficient signal quality for wind profiling at SNR < -6 dB. The uncorrelated noise values show a clear lower bound and are well defined for all values of SNR. The defined range and smooth behavior as a function of SNR enables reliable estimation and removal of the uncorrelated noise for each leg and each measurement altitude individually. The few identifiable outliers above the bounded variance level vanish if a longer minimum time series length is required, they are therefore identified as uncorrelated noise overestimation due to short leg lengths and high atmospheric variance.

Based on this analysis, for the subsequent analysis, the uncorrelated noise is estimated using the linear fit method. Values with uncorrelated noise $< 0.001\,\mathrm{m^2\,s^{-2}}$ are set to $0.001\,\mathrm{m^2\,s^{-2}}$, in line with the above reasoning due to the convex shaped autocovariance functions in the high atmospheric variance regime. The replacement is justified, as negative uncorrelated noise occurs at values of the atmospheric variance $> 0.5\,\mathrm{m^2\,s^{-2}}$, where the variance due to uncorrelated noise becomes negligible ($< 0.03\,\mathrm{m^2\,s^{-2}}$ for SNR > 0 dB).

I am not aware of any method to estimate the random uncertainty $\overline{\varepsilon_{rv}^2}$ associated with the uncorrelated noise estimation. Therefore, a simple approach is chosen here: Based on the above discussion, the random uncertainty is calculated as the average difference between the uncorrelated noise levels estimated using the first lag versus the linear fit estimation method (for the atmospheric variance thresholded points on a 1 dB interval).

Overall, the level of uncorrelated noise is very low even for a lidar measurement frequency of 10 Hz, pointing to an excellent lidar measurement quality. Further, for low levels of atmospheric variance the level of uncorrelated noise can be accurately estimated, whereas for higher levels of atmospheric variance it is negligible. The low levels of uncorrelated noise for all levels of SNR are aided by the viewing geometry of the ADL compared to ground based systems. The ADL system is looking through the free troposphere with less aerosol (lower SNR) down into the BL with more aerosol (higher SNR). Thereby, first the low signal return regime is encountered by the lidar beam, but due to the high signal power sufficient return signal is generated despite the low values of the backscatter coefficient. As the lidar pulse enters and travels through the BL, the signal power rapidly weakens with range, as more aerosol is present which absorbs and scatters the lidar signal along the way. However, due to the increase of aerosol content with distance, sufficient return keeps being generated for high quality measurements throughout the BL.

Scanning system changes

For a scanning system, the lidar signal is expected to be of similar quality due to the similar system setup (a noticeable exception is an upward looking scanner, where ground based signal characteristics apply, leading to reduced signal quality above the BL). A slightly increased level of uncorrelated noise is possible due to transmission losses occurring when the lidar beam passes the mirror as well as the protective window. The level of noise can be estimated using the same analysis procedures as presented here.

6.2.IV. Platform motion correction

After synchronization of the aircraft and lidar data, beam pointing-angle calibration and quality control of the lidar data, the motion correction is applied to all acquired ADL data. In this way, the aircraft contribution to the measured radial velocities is removed and the atmospheric contribution is extracted.

This section outlines the procedure of motion correction and quantifies the associated error based on the measured ground return velocities. The techniques developed as a part of this work allow for assessment of the aircraft motion correction accuracy for a wide range of measurement conditions. It is shown that even disturbed flights inside the BL allow for high accuracy radial velocity measurements. For highest accuracy measurements an improved motion correction using the measured ground return velocities is presented, similar to the procedure proposed in Ellis et al. (2019).

Demonstration of motion correction procedures

The motion correction procedure is illustrated for the first PRY maneuver during the morning flight on 18 July 2019. PRY maneuvers consist of extreme aircraft movements, which exceed the pitch, roll and yaw angles as well as rates encountered during any measurement leg used for meteorological analysis later on (Appendix A.6). PRY maneuvers therefore provide a test environment for the functionality and quality of the motion correction procedure. They present a conservative upper limit of the error expected due to motion correction for normal measurement conditions, including for strongly disturbed flight conditions inside the BL (for a more detailed description see Appendix A.6). A PRY maneuver during a morning situation is chosen because the atmospheric conditions at this time are favorable for accuracy assessment. The vertical profile of the horizontal wind (Fig. C.26) shows moderate wind speeds in the range 2 - 6 $\mathrm{m\,s^{-1}}$ with a maximum at approximately 500 m (400 m AGL). The wind direction is approximately south (around 170°) in the upper part of the profile, changing by more than 45° to approx.

southeast (around 120°) in the lower part of the profile. The aircraft heading during the PRY maneuver is 180°, giving an upwind heading in the upper part of the profile, which changes towards a crosswind heading for the lower part of the profile (Tab. B.3). The terrain in the vicinity of the maneuver is mostly flat (elevation change < 100 m). Due to the early morning situation, the BL is very shallow (O(100 m) AGL), with little turbulence in the residual layer from the previous day and the undisturbed free troposphere above. Consequently, the lidar measurements contain very little turbulence, making an identification of measurement errors due to platform motion correction easier.

Figure 6.8 shows a 2D curtain display of the measurements obtained during the PRY maneuver. The corresponding aircraft roll, pitch and heading angles are shown in Fig. 6.9, alongside with the radial velocity measurements from individual range gates at three different levels and the aircraft vertical velocity measurement. The lidar SNR is sufficient for high quality measurements during the whole PRY maneuver and over the full range, making the contribution of uncorrelated noise small (Fig. 6.8d). The raw radial velocity (Fig. 6.8a) shows the influence of the aircraft maneuvers on the measured radial velocity. During every pitch oscillation the aircraft moves up and down (relative to the surrounding air), which clearly influences the radial velocity. In comparison, the roll movements have a much smaller effect on the measured radial velocity, as less aircraft velocity is introduced into the beam during these maneuvers. The same is applicable for the yaw maneuvers, although these are conducted at much lower amplitude (flight dynamics of the aircraft allow for less deflection in this direction). Further, as the roll maneuvers show the largest amplitude, the beam path through the atmosphere is greatly elongated. The effect is visible through the high-frequency variation of the number of ground range gate during the roll maneuver (resembling grass on the ground, this correlation was used for time synchronization in Sec. 6.2.I).

For motion correction, two procedures have been developed and are analyzed here. First, the raw motion correction is conducted by fully relying on the aircraft navigational data available through the INS (Fig. 6.8b). Results presented in the next section show that the accuracy of this procedure is usually reliable enough for usable measurements, except in turns or during ascent/descent, when INS errors can become large. The advantage of the purely aircraft-based correction procedure is its permanent availability at all times of flight. Second, if a ground return is available, the motion corrected radial velocity can be further corrected for the remaining ground return residual velocity in a refined motion correction procedure. The refined motion correction has been used for ADR by Ellis et al. (2019), and for ADL calibration purposes by Lux et al. (2018). To my knowledge, a refined motion correction at highest temporal resolution has not been applied for ADL before. The procedure uses the 1 s moving averaged ground

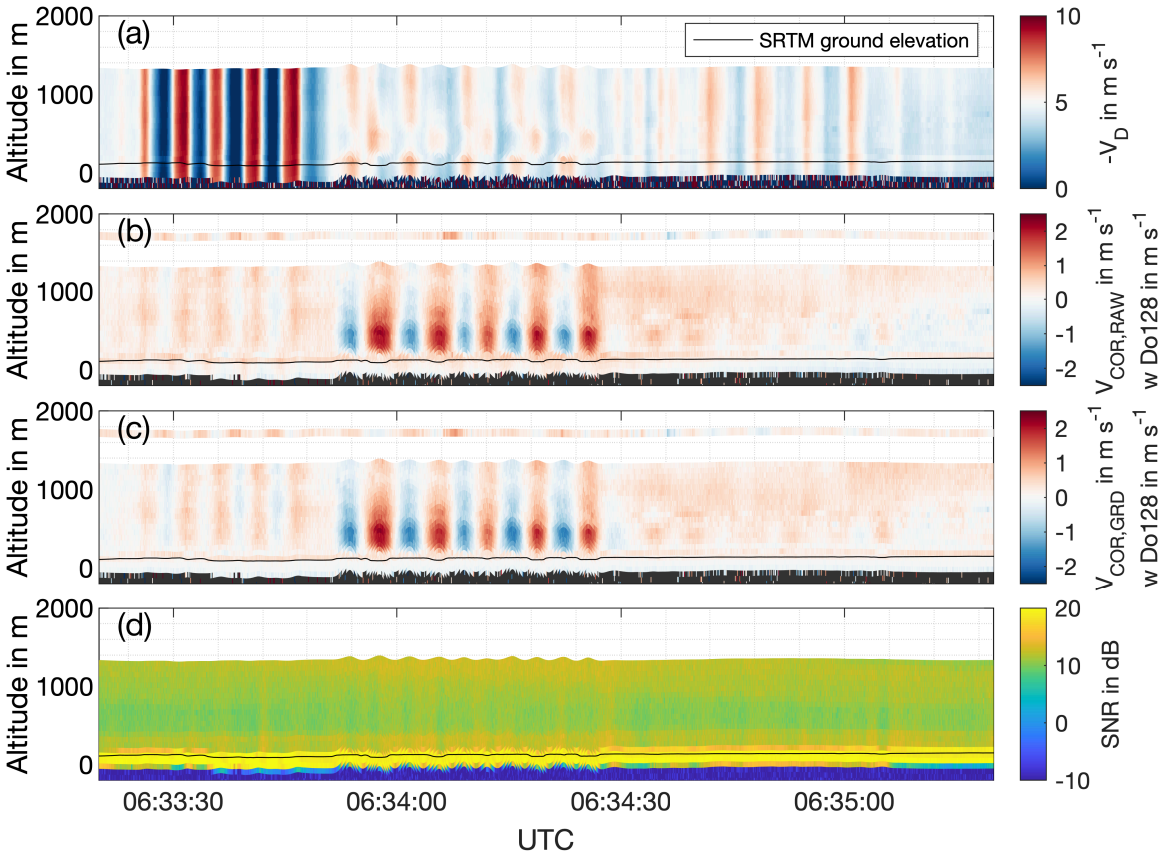

Fig. 6.8.: Illustration of motion correction procedure for PRY maneuver during the morning flight on 18 July 2019. Red colors depict flow towards the lidar, blue colors flow away from the lidar. For simplicity, the measured 3D spatio-temporal data is displayed using a 2D curtain as a function of time. (a) Uncorrected lidar radial velocity measurement. (b) Raw motion corrected lidar radial velocity measurement (atmospheric contribution), obtained from correction using only aircraft INS information (color coded curtain). Additionally, the aircraft vertical wind measurement is extended vertically by ±50 m and displayed at flight level as a color coded line. Please note the adjusted colorbar range. (c) Refined motion corrected lidar radial velocity measurement (atmospheric contribution), additionally corrected using the 1 s moving average of the ground return velocity residual. Aircraft vertical wind measurement as in panel above. (d) Lidar SNR. The ground elevation determined using a DEM is displayed as a black line in all panels.

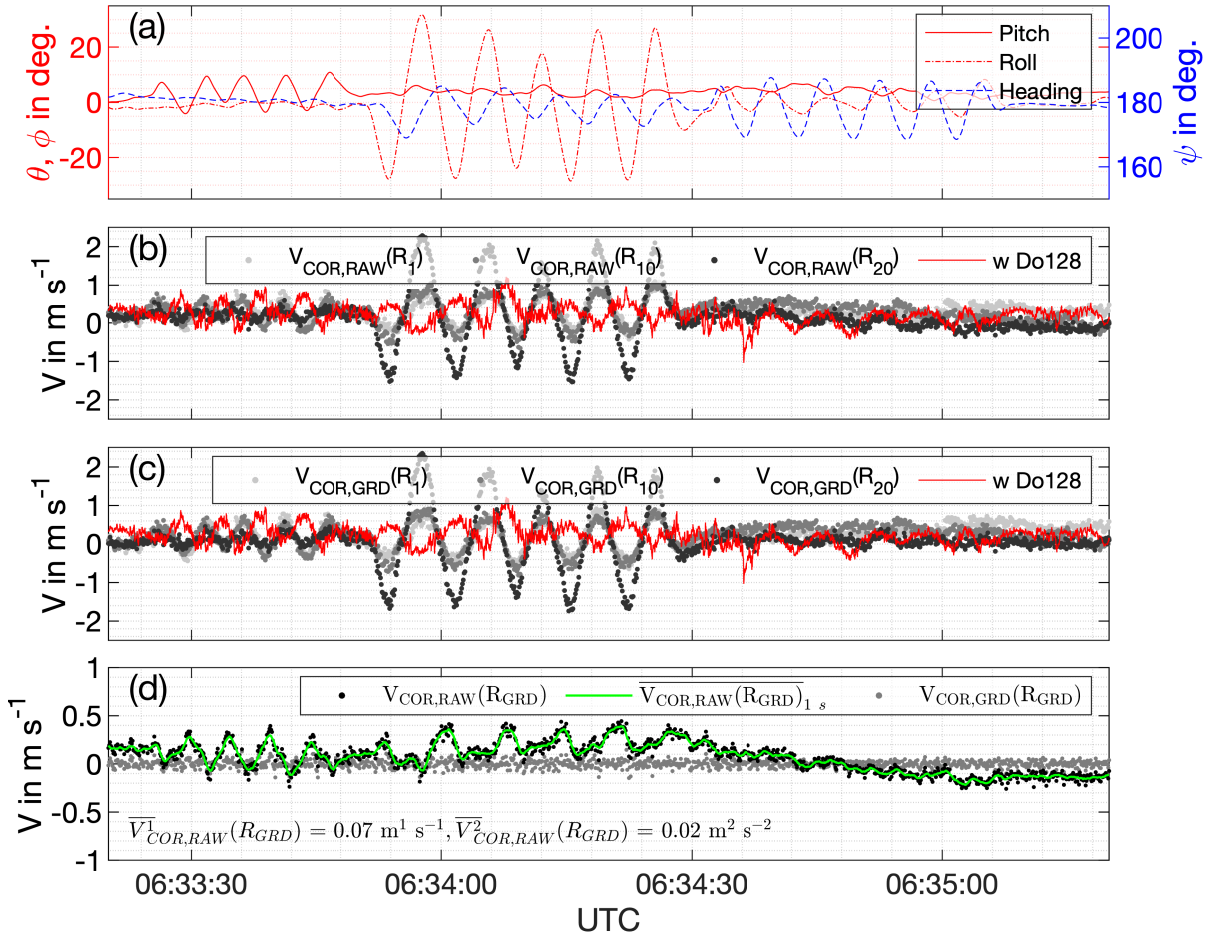

Fig. 6.9.: Data for motion correction procedure for PRY maneuver during the morning flight on 18 July 2019 (Fig. 6.8). (a) Aircraft measured pitch, roll and heading angles. (b) Raw motion corrected lidar radial velocity measurement, obtained from correction using only aircraft INS information, for three different distances from the lidar. Additionally, the aircraft vertical wind measurement is also displayed. Please note that the two quantities are different, as the lidar radial velocity measurement contains the horizontal wind contribution. (c) Refined motion corrected lidar radial velocity measurement, additionally corrected using the 1 s moving average of the ground return velocity residual. Aircraft vertical wind measurement as in panel above. (d) Ground return velocity residual and 1 s moving average used for refined motion correction. Please note the adjusted y-axis limit.

return velocity (Fig. 6.9d), in order to exclude ground return variation with higher frequency. Higher frequencies do not necessarily contain physical information on the motion correction accuracy (e.g. due to moving ground targets such as leaves, crops on the ground, see Fig. 6.9, extensively discussed in Ellis et al., 2019). Using the refined procedure the INS errors can be further corrected, yielding very high accuracy measurements as shown in the following. However, this procedure is only available when usable ground returns are identifiable (e.g. no clouds, which is the case for > 80 % of the measurements, see Appendix A.5 for statistics).
Both procedures remove the majority of the measured radial velocity during the pitch maneuvers, as can be seen in Fig. 6.8b, c. During the roll and yaw maneuvers, again both procedures perform similarly and extract the atmospheric contribution due to the horizontal wind (which does not change noticeably during the yaw maneuvers). The magnitude of the atmospheric contribution varies with altitude due to the changing wind speed and direction, with a maximum below 500 m (Fig. 6.8, 6.9b, c). The discussed maneuver is conducted in an upwind direction for the upper part of the profile, changing to a more crosswind component towards the lower part of the profile. Consequently, at a southerly heading, the pitch maneuver is prone to sample more of the horizontal wind, especially in the upper part of the profile. However, as the pitch amplitude is only 1/3 of the roll amplitude, the dominant effect visible is the atmospheric contribution in the roll maneuver. The roll effect becomes even more dominant towards the lower part of the profile, where the u and v components of the wind are approximately equal (southeasterly direction) and the horizontal wind speed increases to $6\,\mathrm{m\,s^{-1}}$.
While the raw correction yields slightly more positive values compared to the refined correction, it is impossible to say which procedure is more reliable from these results alone, as the atmospheric contribution is still included in the measurements (it is removed in the next section). However, Fig. 6.9d shows that the measured ground return velocity residual is clearly in phase with the pitch as well as roll maneuvers, giving rise to the assumption that there are significant errors present in the INS. Further, Fig. 6.9c shows that the aircraft in-situ measured vertical wind is also noticeably influenced by the PRY maneuvers, which it should not be for an ideal aircraft measurement. Again, the aircraft vertical wind measurement is correlated with the ground return velocity residual (but not exactly in phase), pointing to INS errors as a source. Please note that the aircraft measured vertical wind has to be different from the lidar measured radial velocities as no horizontal contribution is included. The two measures are only comparable after removal of the horizontal wind contribution in the lidar measurement (Sec. 6.2.VI).
Summarizing, it is assumed that the ground return velocity correction is able to identify INS navigational errors. This assumption is confirmed in the following sections, additionally it is shown that the refined motion correction outperforms the raw motion correction considerably.

Motion correction accuracy assessment using measured ground return velocities

The measured ground return velocities can also be used to perform an overall motion correction error assessment and quantification. As discussed in Sec. 3.5, assuming a perfect motion correction procedure and ideal ground return velocity measurement, the ground return velocity should always be $0\,\mathrm{m\,s^{-1}}$ after motion correction. The actual observed ground return velocities include errors due to time synchronization, beam pointing-angle calibration and motion correction as well as ground return velocity measurement (e.g. errors in ground determination, lidar radial velocity noise). The lidar radial velocity noise is not included in this error assessment as the ground return shows systematically higher SNR compared to the atmosphere and thereby incurs less uncorrelated noise. Consequently, an analysis of the ground return velocities allows for a conservative estimation of the error associated with the time synchronization, beam pointing-angle calibration and motion correction procedure.

Ground return mean

Figure 6.10 shows the leg-averaged mean values of the ground return velocity as a function of aircraft pitch variance using the two previously discussed motion correction procedures. Aircraft pitch variance is chosen as the abscissa, as it is closely associated with the vertical acceleration and motion of the aircraft, which is the most difficult component to determine using an INS (Lenschow and Sun, 2007). Further, the pitch angle fluctuates around an almost constant value for individual flight segments (as otherwise stable flight could not be maintained), making the calculation of a variance meaningful. A histogram of the ground return velocities as a function of distance from the ground level, obtained using the raw motion correction procedure, is shown in Fig. A.2.

Starting with the values obtained from nadir legs, the mean ground return velocities cluster around $0\,\mathrm{m\,s^{-1}}$ if the raw motion correction procedure is used, as desired (Fig. 6.10a). The standard deviation of the mean ground return velocity is low ($0.021\,\mathrm{m\,s^{-1}}$) and no correlated increase with pitch variance is observed. For turn legs, the average ground return velocity is also centered around $0\,\mathrm{m\,s^{-1}}$, showing that there is no systematic error in the motion correction. However, the standard deviation is noticeably increased to $0.053\,\mathrm{m\,s^{-1}}$. This observation explains why no reliable beam pointing-angle calibration can be achieved from turn legs. PRY maneuvers show the highest pitch variance, exceeding the pitch variance encountered during any normal measurement leg. Thereby their suitability as a test environment for motion correction accuracy under extreme conditions is confirmed. The ground return velocity measurement

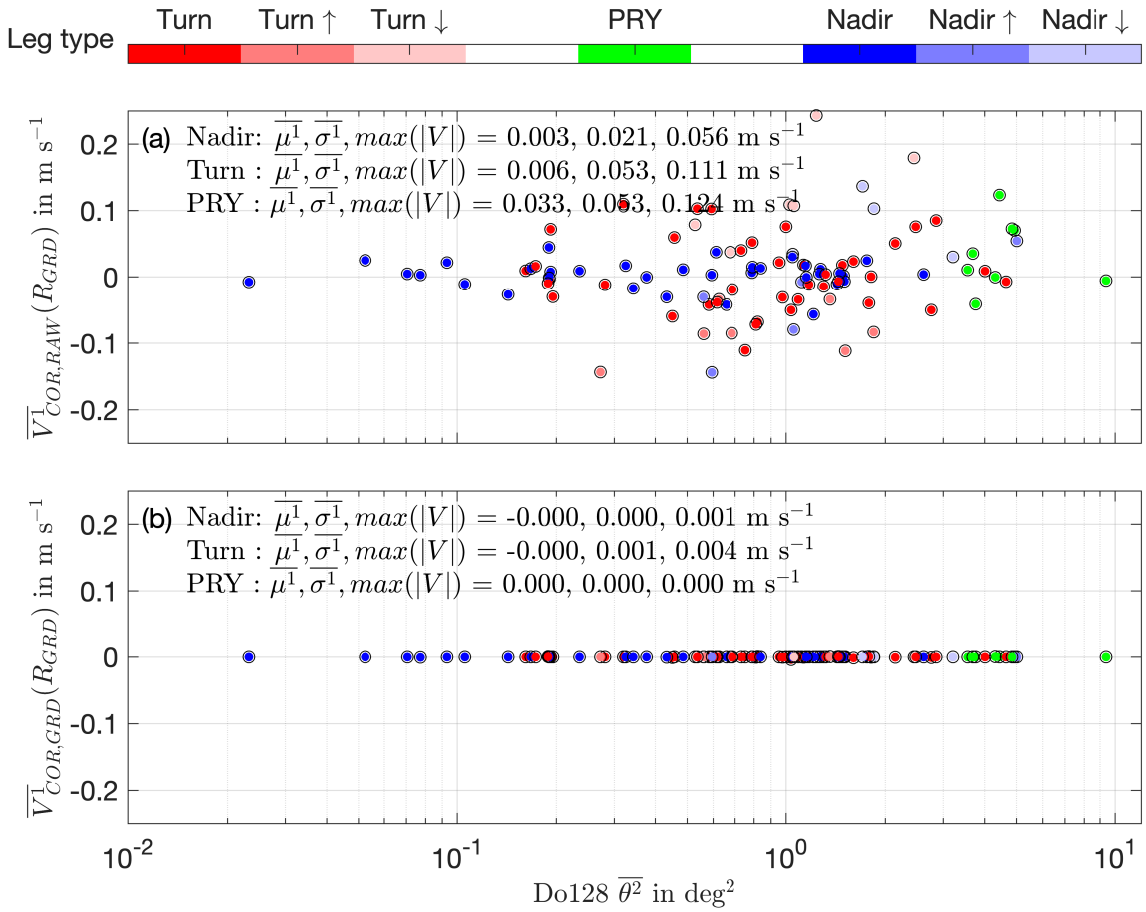

Fig. 6.10.: Ground return based assessment of motion correction errors for mean values using two different motion correction procedures. Displayed is the leg-averaged mean ground return velocity obtained using (a) only the aircraft data for motion correction, and, (b) the refined motion correction procedure including ground return velocity residual correction. Results from all flights and all legs are displayed as a function of aircraft pitch variance. The leg type is color coded.

error for PRY is comparable to that during turns. An interesting feature is observable for ascending versus descending turns and nadir segments. All ascending legs shows systematically low-biased ground return velocities, whereas the opposite is true for descending legs. This observation proves that there is a systematic error in the INS determined vertical aircraft velocity during ascending and descending legs (which also biases the aircraft measured vertical wind, not shown). Consequently, ascending or descending legs should not be used for analysis without the availability of ground return velocity measurements, which can correct this systematic error.

When using the refined motion correction procedure, the mean is brought to 0 m s^{-1} independent of the flight situation, as expected and mandated by the correction algorithm (Fig. 6.10b).

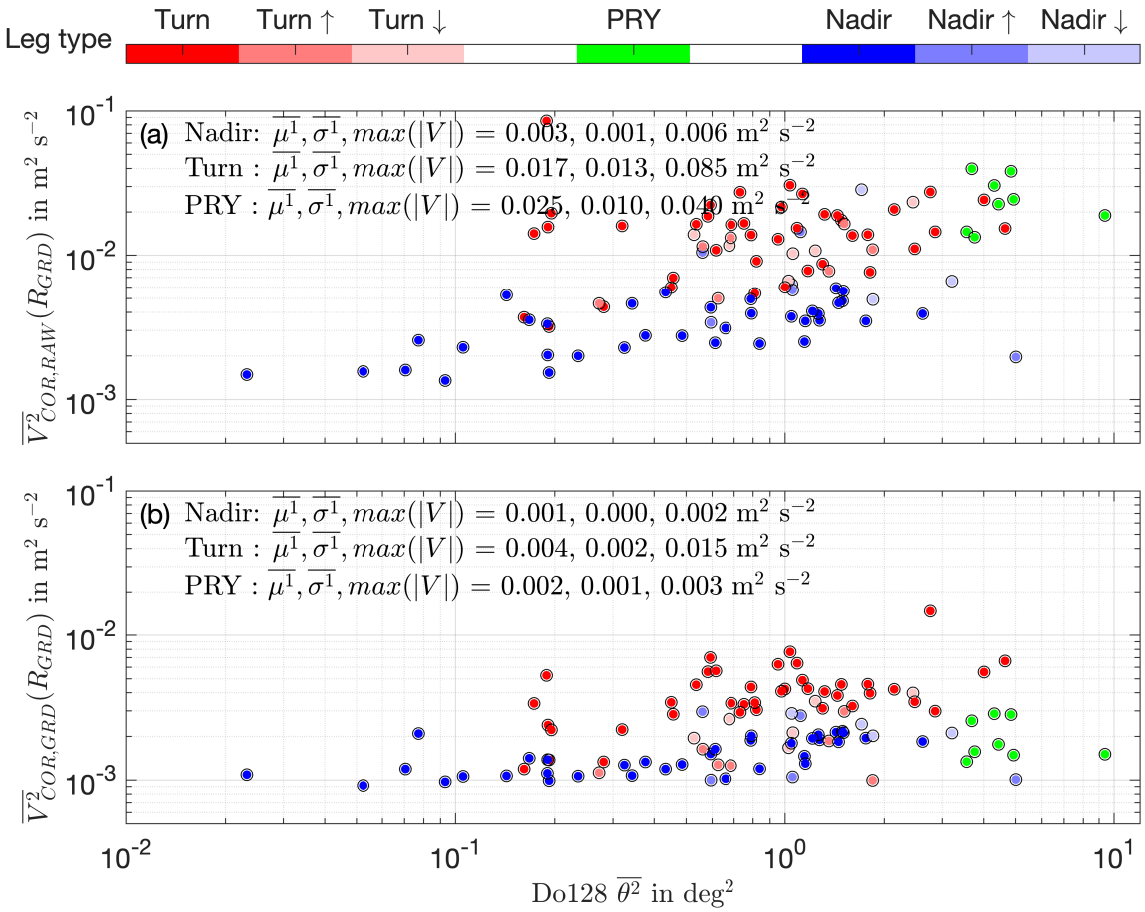

Fig. 6.11.: Ground return based assessment of motion correction error for variance measurements using two different motion correction procedures. Displayed is the variance of the ground return velocity for every leg, obtained using (a) only the aircraft data for motion correction, and, (b) the refined motion correction procedure including ground return velocity residual correction. Results from all flights and all legs are displayed as a function of aircraft pitch variance. The leg type is color coded.

Ground return variance

The measured ground return velocity variance for each leg gives additional insight into motion correction errors (Fig. 6.11). Using the raw motion correction, the measured ground return variance is low for all legs, $\overline{V^2_{COR}}(R_{GRD}) < 0.1\,\mathrm{m^2\,s^{-2}}$. There is a clear separation in ground return variance between nadir and turn legs. Straight and level nadir legs show systematically lower ground return variances ($< 0.01\,\mathrm{m^2\,s^{-2}}$), whereas turn legs show systematically higher ground return variances. This observation is important, as it illustrates that the accuracy of the INS is much higher during nadir legs compared to turns. Both nadir and turn legs show an increase in ground return variance with increasing pitch variance, pointing to an increased raw motion correction error for disturbed flight situations (if no ground return is available for correction).

Nevertheless, the overall variance level for nadir legs remains so low that this increase is not concerning and reliable variance measurements are possible from a motion correction perspective. Caution has to be used when analyzing variance measurements from turn legs without the availability of ground return measurements (an exceptional case in any way). Once again, PRY maneuvers provide an upper bound to the observable ground return variances.
The ground return variance is reduced by almost one order of magnitude ($< 0.01\,\mathrm{m^2\,s^{-2}}$ for turns and $< 0.002\,\mathrm{m^2\,s^{-2}}$ for straight and level nadir legs), as all oscillations longer than 1 s are removed. It is interesting to note that the ground return variance in turns is still elevated in turns compared to nadir legs. Potentially this is due to a more noisy measurement environment in turns, e.g. stronger high frequency vibrations of the aircraft body and increased g-load, which may influence the laser uncorrelated noise level. The previously observed increase in ground return variance with increasing aircraft pitch variance is strongly reduced, demonstrating that the ground return correction works under all flight conditions.
It should be noted that the refined ground return velocity characteristics discussed here, achieved with the refined motion correction, do not prove more accurate measurements by themselves. For example, one could still imagine systematic errors in the ground return velocity determination (e.g. due to chirp, although this effect seems unlikely based on the results in Appendix A.5). In this case, ground return correction would actually worsen the achieved measurement accuracy (which also seems unlikely based on the in-phase variation with aircraft PRY maneuvers, Fig. 6.9). Further, for atmospheric wind measurements, additional effects such as the projection of wind components onto the radial direction need to be accounted for. Therefore, the next section proves that the refined ground return correction actually improves the measurement accuracy and quantifies the improvement due to the refined motion correction procedure based on atmospheric wind measurements.

Summary

Overall, based on the measured ground return velocities a conservative quantification of the errors in the motion corrected radial velocity mean and variance due to time synchronization, beam pointing-angle calibration and motion correction errors, is possible individually for every leg. The average bias of the motion corrected velocities is determined to be negligible for both nadir and turn legs. The standard deviation of the mean is determined as $0.021\,\mathrm{m\,s^{-1}}$ for nadir legs and $0.053\,\mathrm{m\,s^{-1}}$ for turn legs before application of the refined motion correction. Both bias and standard deviation are negligible for measurements obtained using the refined motion correction procedure. The average systematic error of the measured radial velocity variance

due to the aforementioned terms is calculated as the mean of the individual leg variances. It is determined to be $0.003\,\mathrm{m^2\,s^{-2}}$ for nadir legs and $0.017\,\mathrm{m^2\,s^{-2}}$ for turn legs. The additional variance becomes negligible when using the refined motion correction procedure. Theoretically, the additional variance could be removed again from the atmospheric variance for each leg based on the measured ground return variance over the respective leg. However, due to its small magnitude, as well as the possible overestimation due to other factors influencing the ground return velocity, the variance removal is not performed (instead, it is treated as an additional uncertainty).

Scanning system changes

The methodology developed here is applicable to a scanning system in the same way. While the accuracy of the INS is expected to be similarly for a scanning system, a number of additional components can introduce errors in the beam pointing direction and therefore aircraft velocity contribution. Important additional beam pointing errors can be introduced by the movement of the scanner as well as twisting and bending of the whole installation (e.g. movement in the shock mounts, vibration of the scanner boom outside the aircraft, gear play of the scanner mirror). The developed ground return based accuracy assessment and refined motion correction procedure is therefore valuable for performance evaluation of a scanning system as well. Further, the fixed aircraft installation investigated here can provide an important benchmark for comparison as it does not suffer from the discussed additional error sources.

6.2.V. Terminal fall velocity of aerosol scatterers

The combined results of sections (I-IV) show that high accuracy motion corrected radial velocities can be obtained with the prototype ADL system. For wind profiling and nadir measurements of the atmosphere, the radial velocity should be representative for the atmospheric wind contribution. An important factor which could prevent a meaningful analysis is terminal fall velocity of the scattering particles. Their velocity can differ from the atmospheric wind systematically or randomly.

For example, for ADR, the scattering particles are cloud and rain droplets of varying size. Depending on their size, these droplets exhibit a systematic, non-negligible terminal fall velocity which needs to be accounted for. For the used prototype ADL (wavelength $1.6\,\mu$m) the scattering particle size is typically in the range $0.1 - 10\ \mu$m (Mie-scattering regime). Consequently, the terminal fall velocities are on the order of $10^{-7} - 10^{-3}\ \mathrm{m\,s^{-1}}$ ($0.1 - 1000\ \mathrm{cm\,h^{-1}}$), which are negligible for wind measurement purposes (Fig. 6.12, extracted from Seinfeld and Pandis,

2006). Velocity retrieval from clouds or rain must be avoided, as there terminal fall velocity can be larger and an additional chirp effect is present due to the hard target characteristics. The procedures to avoid cloud retrievals are outlined in Sec. 6.2.III.

Random movement direction and speed of the scattering particles is a factor which needs to be considered for ADL. All particles are subject to random motion which is superimposed on the mean motion. The kinetic energy of the random motion is proportional to the temperature of the particles. The scattering aerosols are rather large compared to molecular components, therefore their random motion is smaller. Aerosol motion can be considered as a random process (as the movement of the particles is not systematically oriented), which leads to a broadening of the Doppler spectrum (Werner, 2005). The mean wind is unaffected, as the aerosols do not exhibit a systematic motion. Thereby, aerosol movement within the probe volume also contributes to the radial velocity noise ε_{rv} (Frehlich, 1997). Consequently, the effects of random aerosol movement are generally small and captured as part of the uncorrelated noise estimation. Hence, they do not require separate detailed treatment here.

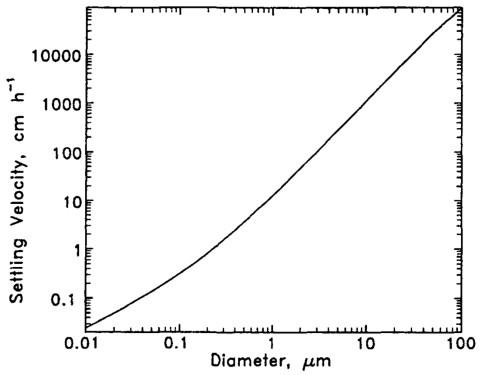

Fig. 6.12.: Terminal fall velocity of particles in air at 298 K and 1000 hPa as a function of particle diameter. Figure taken from Seinfeld and Pandis (2006). © John Wiley and Sons, re-used with permission.

6.2.VI. Non-nadir beam pointing - horizontal wind contribution removal

The combined results of sections (I-V) show that high accuracy motion corrected radial velocities, representative for the atmospheric wind projection onto the beam, can be obtained with the prototype ADL system. Thus, the previous sections are sufficient to demonstrate the reliability and accuracy of the ADL prototype measurement system for wind profiling purposes. For wind profiling, the additional error due to a violation of the flow homogeneity assumption needs to be considered. However, this error is not part of the measurement system accuracy evaluated here as it is investigated using the ADLS in Chap. 5.

Besides wind profiling, an important task of an ADL system for BL research is the analysis of the vertical wind by means of nadir measurements. For this purpose, additional error contributions exist, which need to be further analyzed and quantified (see overview in Sec. 3.5). This

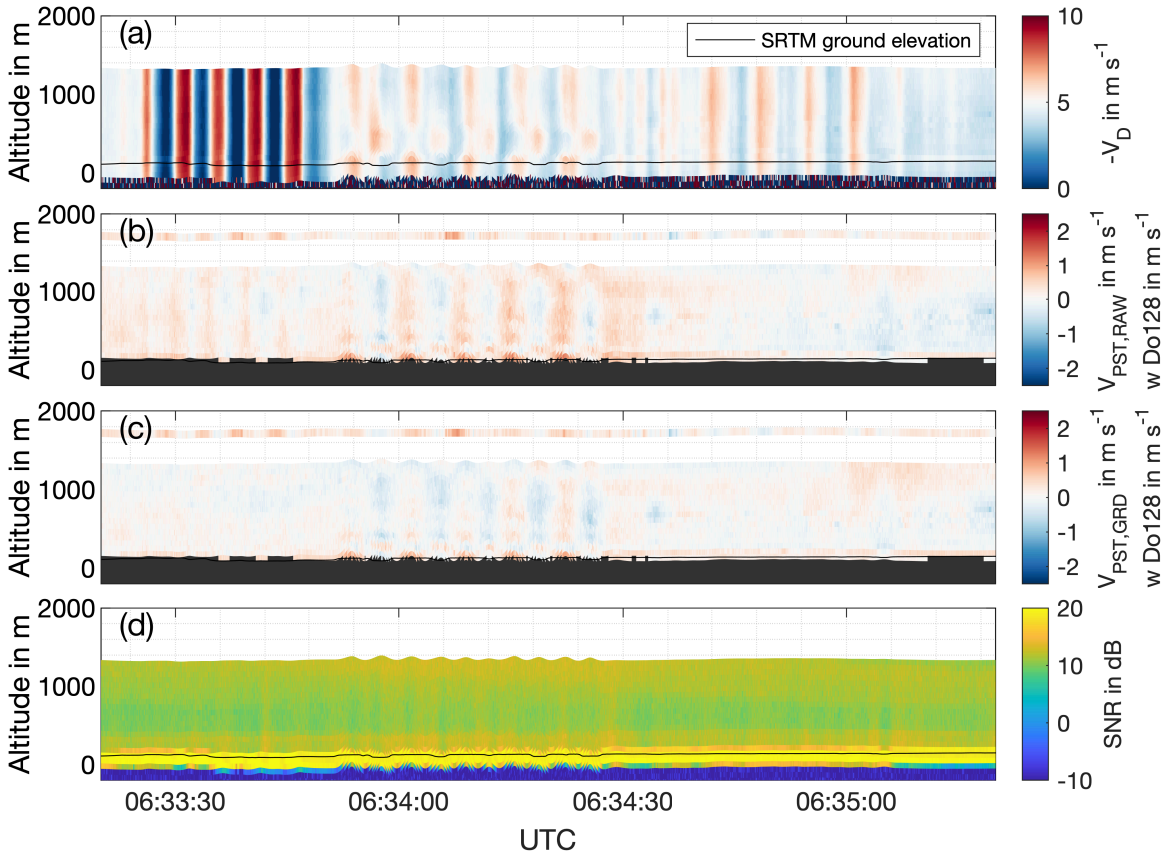

Fig. 6.13.: Illustration of motion and horizontal wind correction procedure for PRY maneuver during the morning flight on 18 July 2019, display characteristics as in Fig. 6.8. As motion and horizontal wind correction is performed, the corrected velocities theoretically show the atmospheric vertical wind contribution to the measurement. (a) Uncorrected lidar radial velocity measurement. (b) Raw motion and horizontal wind corrected lidar measurement (color coded curtain), with motion correction obtained using only aircraft information. Additionally, the aircraft vertical wind measurement is extended vertically by ±50 m and displayed at flight level as a color coded line. Please note the adjusted colorbar range. (c) Refined motion and horizontal wind corrected lidar measurement (color coded curtain). Motion correction obtained by additionally using the 1 s moving average of the ground return velocity residual. Aircraft vertical wind measurement as in panel above. (d) Lidar SNR. The ground elevation determined using a DEM is displayed as a black line in all panels.

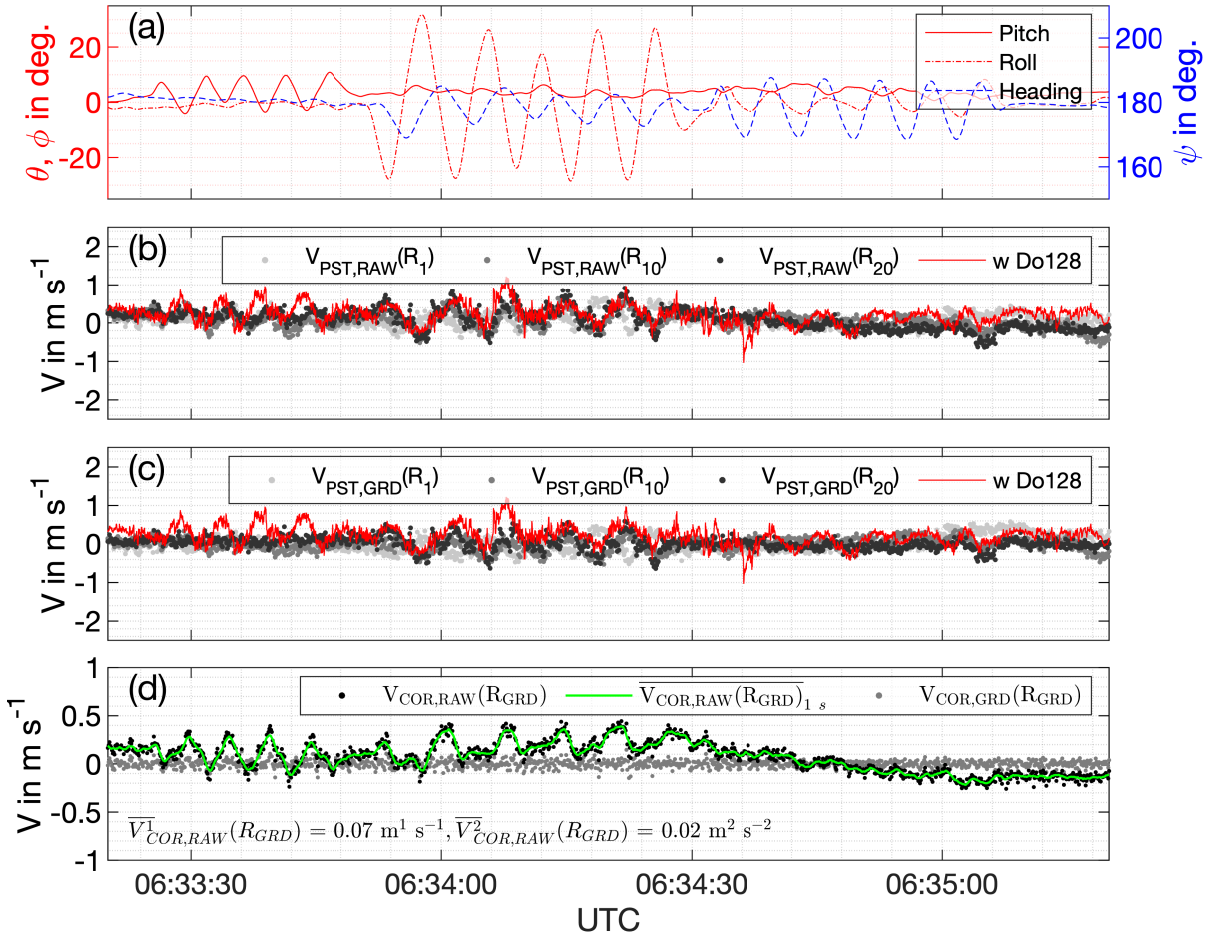

Fig. 6.14.: Data for motion correction procedure for PRY maneuver during the morning flight on 18 July 2019 (Fig. 6.13). (a) Aircraft measured pitch, roll and heading angles. (b) Raw motion and horizontal wind corrected lidar measurement, using only aircraft information, for three different distances from the lidar. Additionally, the aircraft vertical wind measurement is also displayed. Please note that the measurements are taken at different locations both horizontally and vertically. (c) Refined motion and horizontal wind corrected lidar measurement, additionally corrected using the 1 s moving average of the ground return velocity residual. Aircraft vertical wind measurement as in panel above. (d) Ground return velocity residual and 1 s moving average used for refined motion correction. Please note the adjusted y-axis limit.

section deals with the problem that the beam does not point exactly towards nadir at all times. Due to the fixed installation of the lidar inside the cabin, aircraft movements due to pitch, roll and yaw are not compensated. Consequently, the lidar beam deviates from the nadir direction and samples a contribution of the horizontal wind, as previously seen, which is unwanted for vertical wind measurements. The magnitude of the horizontal wind projection depends on the deviation of the beam from nadir as well as the vertical profile of the horizontal wind. If the vertical profile of the horizontal wind is known, this unwanted contribution due to the horizontal wind can be removed again, as done by Leon et al. (2006) and Chouza et al. (2016b).

Demonstration of horizontal wind contribution removal

In the following, the principle of the mean horizontal wind contribution removal (also termed post-processing in the following) is illustrated based on the same PRY maneuver analyzed in the previous section. The horizontal wind present during the PRY maneuver is estimated from the obtained ADL wind profile directly prior to the PRY maneuver (Fig. C.26). Results are presented in Fig. 6.13 and Fig. 6.14 using the previous display arrangement, but showing the lidar measurements corrected for the horizontal wind contribution (V_{PST}, post-processed). Both, the post-processed, raw motion corrected measurements (Fig. 6.13b) and the post-processed, refined motion corrected measurements (Fig. 6.13c) remove the majority of the atmospheric contribution due to the horizontal wind present in Fig. 6.8b, c. However, especially during the roll maneuver, structures in-phase with the roll movement are still identifiable. For the presented PRY maneuver the refined motion correction performs noticeably better, with the measurements centered around the expected $0\,\mathrm{m\,s^{-1}}$ and weaker discernible residual structures. This improved performance is shown to be systematic in the following.

Detailed analysis of the post-processed measurements reveals further interesting features (Fig. 6.14b, c). The raw motion corrected, post-processed lidar measurement still exhibits unwanted noticeable structure during the pitch maneuver, similar to the vertical wind measured in-situ by the aircraft. The refined ground motion corrected lidar measurement, on the other hand, shows no such artifacts and clearly outperforms the in-situ vertical wind measurements during pitch maneuvers. This invariance to pitch maneuvers is an important finding, as it provides confidence in the applied ground correction algorithm even under extreme measurement conditions. The finding also confirms that the vertical aircraft velocity component is the most difficult to estimate using the INS, as potential errors are unbounded (Lenschow and Sun, 2007). For the roll and yaw maneuvers the refined values are centered more around zero, but amplitude-wise both the raw and refined procedure provide similar results, which are also comparable to the aircraft

vertical wind measurement uncertainty. The larger lidar uncertainty during roll maneuvers is founded in the less certain lidar beam pointing-angle in the b_y-direction (Sec. 6.2.II), as well as the smaller aircraft velocities introduced in this direction, making the impact of noise and other measurement errors more important.
As a summary, the results of the horizontal wind contribution removal in post-processing are promising. This is especially true when bearing in mind the extreme non-steady aircraft states reached during the PRY maneuver, which strongly exceed normal measurement conditions. The absolute accuracy of the lidar vertical wind measurement can equal or surpass that of the in-situ aircraft measurement. The additional variance introduced due to the projection of the horizontal wind into the measurement is strongly reduced.

Horizontal wind post-processing accuracy assessment

For overall quantification of the vertical wind measurement accuracy under a range of conditions, this work employs a two-step approach to account for the influence of the horizontal wind projection on both the vertical wind mean and variance.
As a first step, already demonstrated, the contaminating contribution of the horizontal wind to the nadir measurement is strongly reduced by removing the contribution of an assumed horizontal wind field along the flight track (Leon et al., 2006; Chouza et al., 2016b). The horizontal wind field is calculated from quality controlled wind profiles obtained from wind profiling, which was conducted in the vicinity of each nadir leg. For short legs with < 5 min duration the closest available wind profile is used, whereas for longer legs a linearly interpolated wind field is calculated from the preceding and subsequent wind profiles. A well-suited measure to assess the overall error of the post-processed data is the mean vertical wind measured by the ADL over extended legs. For continuity reasons the vertical wind must average to $0\,\mathrm{m\,s^{-1}}$, if longer flight legs are used and no large scale processes (e.g. synoptic forcing, meso-scale phenomena, complex terrain) are present. As a second step, the uncertainty of the calculated vertical wind variance, due to the non-nadir pointing beam, is estimated based on an uncertainty estimation method developed by Strauss et al. (2015). Last, their method is modified to account for the reduced uncertainty due to horizontal wind post-processing in the work presented here.

Vertical wind mean retrieval error

Results for the leg-averaged mean vertical wind measurements are shown for all legs on all flights in Fig. 6.15 for the second range gate and in Fig. 6.16 for all range gates. All measurements with a minimum data availability of 60 s are analyzed.

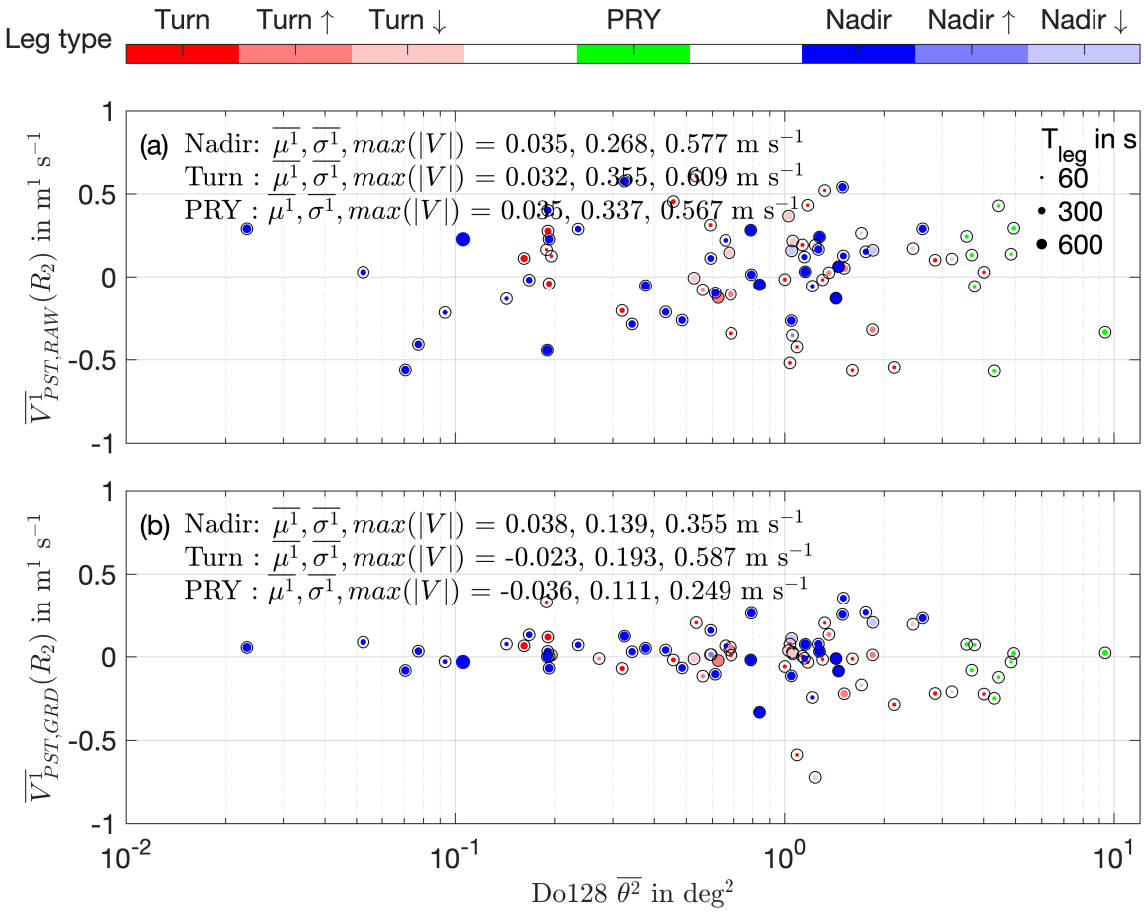

Fig. 6.15.: Leg-averaged mean values of the motion and horizontal wind corrected ADL vertical wind measurements as a function of aircraft pitch variance. In (a) the vertical wind measurements obtained using only the aircraft INS data for motion correction are shown. In (b) the vertical wind measurements obtained using the refined motion correction procedure are shown. Results are shown for all flights and all legs with a minimum length of 60 s. The leg type is color coded, the leg length is displayed through the filled circle area. The displayed data is calculated using the second range gate from the aircraft, as the first range gate contains fewer measurements due to the lidar blind zone proximity.

When using the raw motion corrected data for post-processing (Fig. 6.15a, 6.16a), the values fluctuate around the desired 0 m s^{-1}, but exhibit non-negligible scatter independent of the aircraft pitch variance. The scatter is larger for turns compared to nadir legs (besides the obvious result of larger scatter with shorter averaging), as during turns the motion correction accuracy is compromised (see Sec. 6.2.IV). The scatter present in the nadir leg estimates of the mean vertical wind appears too large to be explainable by large-scale processes, pointing to measurement errors as a cause.

This assumption is confirmed by an investigation of the mean vertical wind obtained from the refined motion corrected data with horizontal wind post-processing applied (Fig. 6.15b, 6.16b).

All flight situations show a much more defined distribution of the vertical wind with strongly reduced scatter, while still being narrowly centered on a mean value of $0\,\mathrm{m\,s^{-1}}$. Both nadir and turn legs show a slight increase in scatter with increasing aircraft pitch variance, which is explained with the difficulty of an accurate horizontal wind projection removal under turbulent conditions (the ground return variance only showed a minor variance increase with increasing pitch variance, therefore decreased motion correction accuracy is unlikely as a cause). Turn legs still show an elevated level of scatter compared to nadir legs, which can be explained by three reasons. First, turn legs are systematically shorter than nadir legs, leading to a less fulfilled zero-mean vertical wind assumption. Second, turn legs are conducted in a more limited spatial area (due to the circle flight path) and thereby also suffer more from local flow phenomena, violating the zero-mean vertical wind assumption. Third, as previously discussed, the refined motion correction accuracy is slightly degraded in turns compared to nadir legs. Nadir legs show very little scatter, therefore the observed residual scatter may well be due to actual meteorological meso-scale phenomena violating the zero-mean vertical wind assumption (see Sec. 6.7 for an example).

For the overall accuracy assessment, the remaining average bias and standard deviation of the mean vertical wind for all range gates on all legs is used for the estimation of an upper error bound. This is clearly a very conservative estimate, since all remaining atmospheric wind contributions present in the measurements are attributed to ADL measurement error. Further, all error sources from the terms (I-V) are included in this estimate as well. Consequently, one obtains a bias of $0.029\,\mathrm{m\,s^{-1}}$ and a random error of $0.219\,\mathrm{m\,s^{-1}}$ for nadir measurements of the mean vertical wind using INS-based motion correction. Using the refined motion correction this accuracy can be improved by lowering the bias slightly to $0.027\,\mathrm{m\,s^{-1}}$ and the random error substantially to $0.125\,\mathrm{m\,s^{-1}}$. Turn legs show slightly degraded accuracy due to the three factors discussed above, but nevertheless very reliable results.

Overall, this analysis shows that the ADL measurements of the mean vertical wind provide excellent accuracy under a wide range of conditions. The achieved accuracy confirms both the reliability of the motion correction procedure as well as the horizontal wind removal in post-processing. The refined motion corrected data performs superior to the raw motion corrected data. This outcome was suggested by the results in the previous section already and the mean vertical wind analysis provides the final confirmation. Although the raw motion corrected data is far from being unusable, the added value of the refined motion correction procedure is substantial. Consequently, in the following all presented results are obtained using the refined motion correction procedure, whenever a usable ground return signal is available.

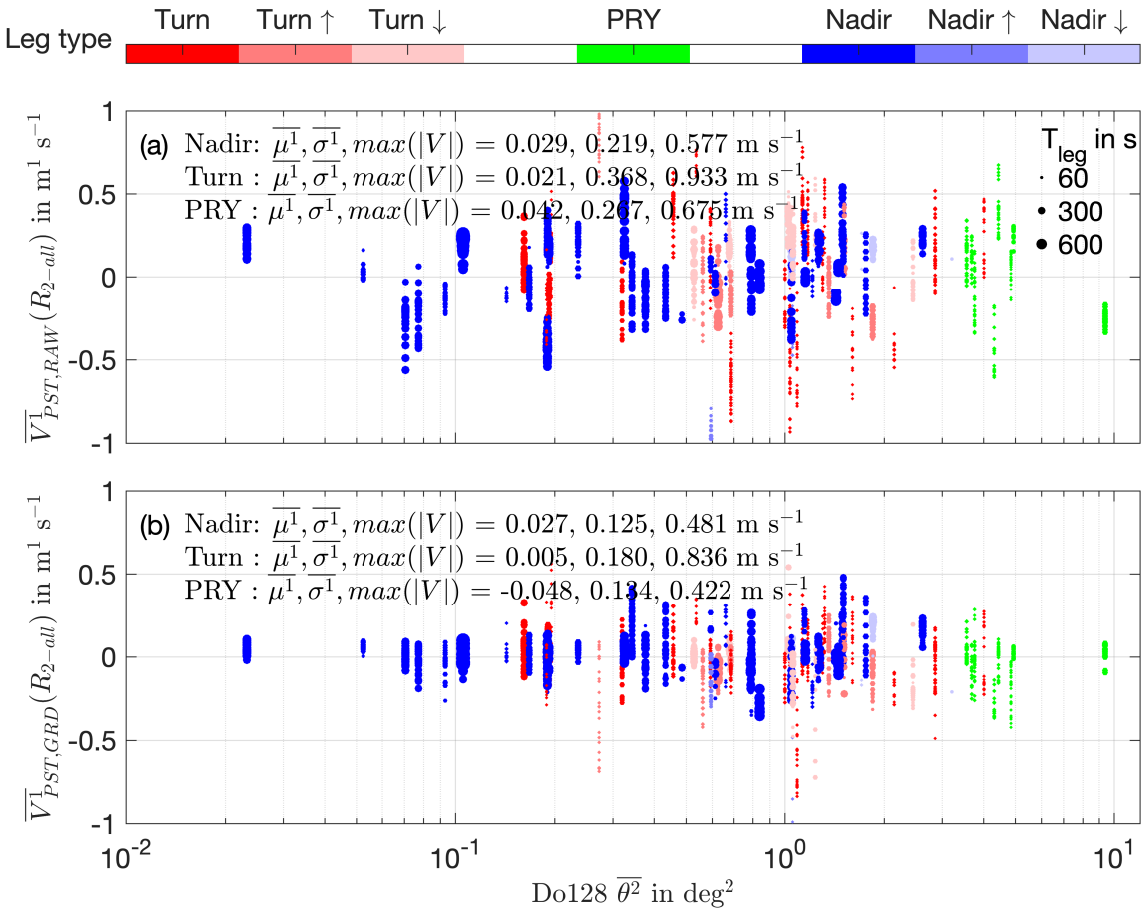

Fig. 6.16.: Leg-averaged mean values of the lidar motion and horizontal wind corrected vertical wind measurements as a function of aircraft pitch variance. In (a) the vertical wind measurements obtained using only the aircraft INS data for motion correction are shown. In (b) the vertical wind measurements obtained using the refined motion correction procedure are shown. Results are shown for all flights and all legs with a minimum length of 60 s. The leg type is color coded, the leg length is displayed through the filled circle area. The displayed data is calculated using all range gates starting from the second.

As a side note, one can also compare the leg-averaged lidar data to the leg-averaged in-situ aircraft measurements for investigation of measurement accuracy, although the two measurements are collected at different altitudes and therefore may suffer from systematic differences. For nadir legs, this comparison yields results similar to the pure lidar results discussed here. However, the aircraft vertical wind measurement is strongly positive biased for left turns, and slightly negative biased for right turns, making the comparison meaningless in turns. The biased aircraft measurements during turns were discussed with the operator of the aircraft but no solution could be found before the completion of this work (see previous discussion on INS accuracy and Sec. 6.3).

Accuracy of the vertical wind variance retrieval

The variance obtained using the post-processed, horizontal wind corrected data cannot be easily evaluated for uncertainty as is possible for the mean, because it is not bounded by a constraint similar to the zero-mean vertical wind. Direct comparison to the in-situ aircraft measurements is also not advisable, as the variance exhibits a non-constant vertical profile and other factors such as pulse-volume averaging also play a role (Sec. 6.2.VII).

Due to the difficulty in assessing the influence of the horizontal wind projection on the measured variance directly (see Fig. 6.13 for an example with extreme aircraft movement), an extensive method for uncertainty estimation has been developed by Strauss et al. (2015). The method was developed for usage with raw motion corrected data (not corrected for horizontal wind contribution in post-processing), as the assumption of a linear wind field is not fulfilled in their study due to measurements in complex terrain. Therefore, the method is adapted here to provide uncertainty estimates when using horizontal wind contribution post-processed data (which is preferable as the removal of the horizontal wind contribution greatly reduces the erroneously introduced variance).

The uncertainty estimation method developed by Strauss et al. (2015) splits the erroneous variance contribution in the lidar signal into two separate factors (see Appendix A.7 for the theory and formulas associated with this paragraph). First, additional variance $\overline{\beta_{HC}^2}$ is contributed by the horizontal wind projection, e.g. due to variance in aircraft pitch and roll, creating a biased variance measurement as this contribution can only be positive. In their study, the additional variance is created due to the atmospheric mean wind contribution (Fig. 6.9c), as well as due to the additional movement of the lidar beam along the track (e.g. slicing through additional up- and downdraft to the left and right of the track due to roll maneuvers). Their uncertainty estimation only provides an upper bound of the possible systematic offset and not an exact estimate, therefore, the systematic offset cannot be removed (as is done for the random radial velocity noise contribution) but has to be treated as an additional systematic uncertainty. In this work, the former contribution is strongly reduced using the discussed post-processing of the lidar measurements to remove the horizontal wind contribution. The latter term for additional variance cannot be mitigated when using a fixed beam system and needs to be accounted for (it could be mitigated when using a motion compensated scanning lidar system). Second, an additional random uncertainty $\overline{\varepsilon_{HC}^2}$, which can have either sign, can be present in the measurements, e.g. due to non-zero mean pitch and roll angles (there is a constant pitch-up angle for the Do128 for regular straight and level flight, increasing with flight altitude) and other contributions.

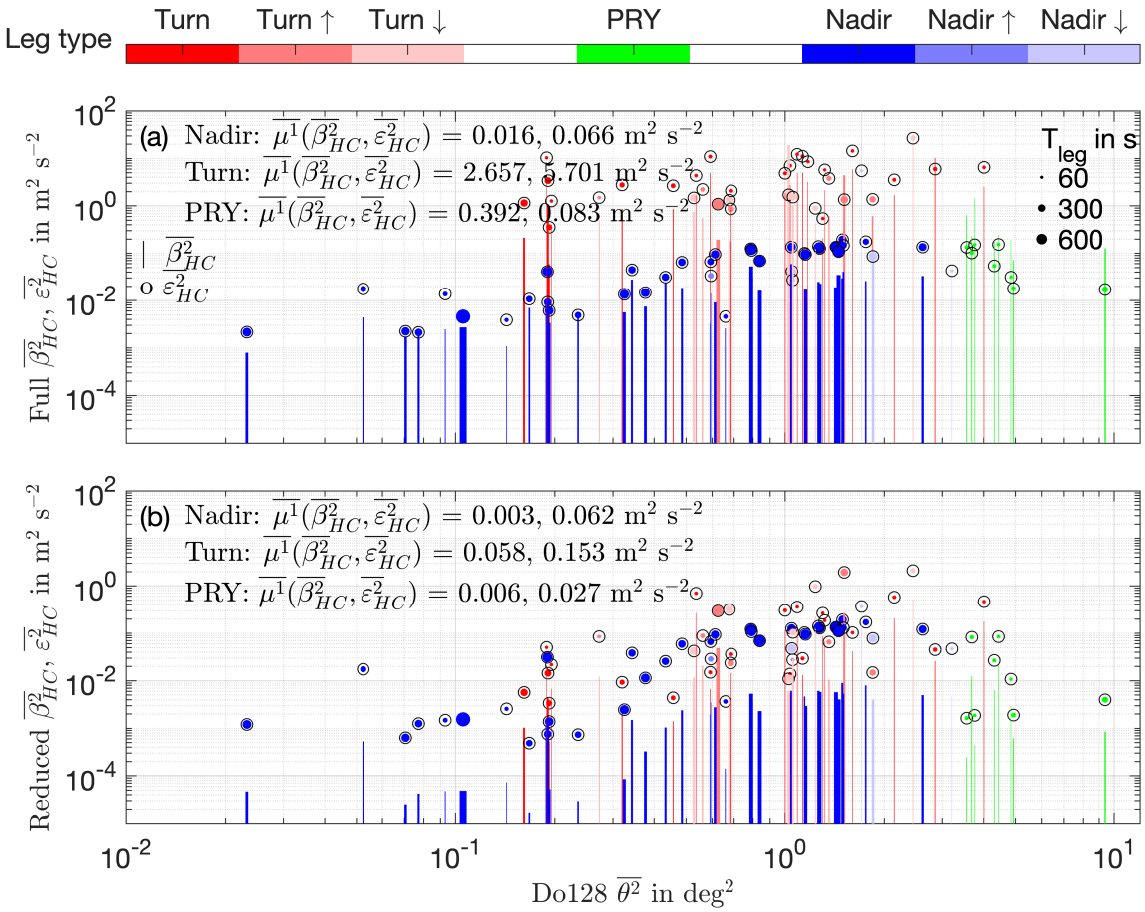

Fig. 6.17.: Uncertainty of the lidar measured vertical wind variance due to horizontal wind contamination as a function of aircraft pitch variance, calculated based on Strauss et al. (2015). Vertical lines show the range up to the maximum positive systematic offset $\overline{\beta^2_{HC}}$ of the measured variance. Circles show the random uncertainty $\overline{\varepsilon^2_{HC}}$ in the measured variance, which can be a negative or positive contribution. (a) shows both terms considering the full horizontal wind projection as a contaminating source, assuming no horizontal wind correction in post-processing (as in Strauss et al., 2015). (b) shows the reduced version of both terms, which can be achieved if a linear wind field is assumed and horizontal wind correction is applied in post-processing. Results are shown for all flights and all legs with a minimum length of 60 s. The leg type is color coded, the leg length is displayed through the filled circle area and line width respectively. The displayed data is calculated using the second range gate from the aircraft, as the first range gate contains fewer measurements due to the lidar blind zone proximity. Please note that (differing from previous figures) both panels use the more accurate refined motion corrected data.

Both uncertainty contributions are visualized in Fig. 6.17. The full uncertainty estimates based on Strauss et al. (2015), using the refined motion corrected data but without horizontal wind removal in post-processing, is shown in Fig. 6.17a. As expected, variance measurements from turns legs without removal of the mean wind contribution are impossible, as the main variance contribution is due to the mean wind signal. Consequently, turn legs suffer from high possible

systematic offsets due to non-zero mean pitch and roll ($\overline{\beta_{HC}^2}$). Turn legs also have increased random uncertainty due to $\overline{\varepsilon_{HC}^2}$, which is also expected as pitch and roll are more variable during turns compared to nadir legs. For nadir legs, there is a general increase of the uncertainty with increasing pitch variance for both $\overline{\beta_{HC}^2}$ and $\overline{\varepsilon_{HC}^2}$, as both contain the sampled radial velocity variance as important terms, which increases with increasing pitch variance. The potential systematic offset $\overline{\beta_{HC}^2}$ increases less compared to the random uncertainty $\overline{\varepsilon_{HC}^2}$. Compared to the previously discussed uncertainties, the levels observed here are non-negligible ($\overline{\varepsilon_{HC}^2}$ up to $0.2\,\mathrm{m^2\,s^{-2}}$, average $0.066\,\mathrm{m^2\,s^{-2}}$ for nadir measurements) and can present an important contribution to the lidar measured variance.

Due to the potentially high magnitude, a reduction of the discussed uncertainty is desirable for ADL measurements inside the disturbed BL. To this end, the theory presented by Strauss et al. (2015) is combined with the procedure of Chouza et al. (2016b) for horizontal wind contribution removal here, resulting in a modified version of the Strauss et al. (2015) theory, explained in Appendix A.7.

Figure 6.17b shows the reduced uncertainty estimate obtained using the new approach. As expected, both the systematic offset and random uncertainty during turn legs are greatly reduced. Thus, variance measurements are also possible during turn legs (as conducted by Lothon et al., 2005, for ADR already, however, without estimating the uncertainties discussed here). Even better accuracy can be achieved for nadir legs, where the potential systematic offset is further reduced to negligible values $< 0.01\,\mathrm{m^2\,s^{-2}}$. The random uncertainty is also reduced, although less, and a noticeable increase with aircraft pitch variance up to values of $0.15\,\mathrm{m^2\,s^{-2}}$ remains. Overall, the random uncertainty in nadir variance measurements due to $\overline{\varepsilon_{HC}^2}$ is not generally negligible and can cause noticeable uncertainty for disturbed measurement flight conditions. Fortunately, a leg-specific, conservative estimation is possible through the modified Strauss et al. (2015) framework.

Scanning system changes

For a scanning system, the above discussed errors can be mitigated or at least strongly reduced, if the scanner is able to compensate for the aircraft attitude in real time, based on transmitted INS data. In this case, the beam points towards nadir at all times, and thereby no contaminating horizontal wind projection is introduced. The limitation in this case is the accuracy of the transmitted aircraft attitude angles (pitch, roll, yaw) as well as the scanner motion compensation speed and accuracy. The theory and algorithms developed here, especially using PRY maneu-

vers, allow for quality control of the motion compensation accuracy of the scanner in the same way as conducted here for the fixed beam installation.

6.2.VII. Pulse-volume averaging

The spatial extent of the lidar pulse and range gate volume cause smoothing of fine-scale structures present in the wind field (Sec. 3.1). In general, this PVA results in an underestimation of the variance and an overestimation of the length scales measured by the lidar (Fig. 3.1), compared to what would be measured by an ideal in-situ measurement (Frehlich, 1997; Frehlich et al., 1998; Frehlich and Cornman, 2002; Davies et al., 2004; Frehlich et al., 2006; Lothon et al., 2006; Stawiarski, 2014; Brugger et al., 2016). To resolve this problem procedures exist to correct for the PVA based on theoretical turbulence models. In this work, PVA is corrected based on a method introduced by Frehlich et al. (2006), which is presented, modified for airborne application and evaluated in the following.

As the moving lidar beam from the ADL samples a two-dimensional plane in the atmosphere, two averaging processes can occur. First and most importantly, along-beam averaging takes places due to the spatial extent of the lidar pulse along the beam. Second, for a nadir staring beam, moving with the speed of the aircraft platform, averaging in the direction of the horizontal movement (orthogonal to the lidar beam direction) also occurs as 75 pulses are combined to produce one radial velocity measurement (duration 0.1 s). Frehlich and Cornman (2002) show that the horizontal averaging, orthogonal to the lidar beam direction, can be neglected if the movement Δh of the lidar during one velocity estimation is much smaller than the range gate length of each velocity estimate ($\Delta h << \Delta p$). In this work, $\Delta h \approx 6.5\,\mathrm{m}$ (10 Hz measurement frequency at an aircraft speed of $\approx 65\,\mathrm{m\,s^{-1}}$) and $\Delta p = 72\,\mathrm{m}$, thereby $\Delta h / \Delta p \approx 0.1$, which is sufficient to fulfill this condition according to Frehlich and Cornman (2002). Consequently, spatial averaging in the horizontal (orthogonal to the lidar beam direction) can be neglected in this work. Thereby, the averaging in along-beam direction remains as an important driver of PVA and needs to be accounted for. In this work, the along-beam PVA is corrected using the azimuthal structure function technique introduced by Frehlich et al. (2006). The azimuthal structure function technique corrects along-beam PVA based on a structure function calculation orthogonal to the beam direction, which should not be confused with the neglect of averaging in this direction. The technique is introduced in the following, it is modified for usage with the vertical wind, its applicability is checked and the uncertainty associated with the correction is evaluated. For the sake of brevity, the associated formulas are not shown here but provided in Appendix A.8 instead.

Pulse volume averaging correction procedure

Due to the lidar viewing geometry, most often the PVA effect is corrected through structure function calculation and analysis in along-beam direction (Frehlich, 1997; Frehlich et al., 1998; Frehlich, 2001; Frehlich and Cornman, 2002). However, a disadvantage of this technique is its coarse vertical resolution for nadir measurements. A number of range gates are necessary to determine the structure function, thus only one PVA corrected value of variance and integral length scale can be obtained for the whole BL. Due to this disadvantage, an extension of existing techniques is proposed by Frehlich et al. (2006). He provides a model of the structure function and associated PVA correction for a two-dimensional plane, calculated for movement orthogonal to the beam pointing direction (while still neglecting spatial averaging in the movement direction, orthogonal to the beam pointing direction, and requiring $\Delta h << \Delta R$, with ΔR being the effective range resolution, $\Delta R = \Delta p + \Delta r$). Consequently, the resolution in along-beam direction is significantly improved, enabling PVA correction for every individual range gate along the beam. Frehlich et al. (2006) proposes a version of the pulse-volume averaging correction for scans over varying azimuths at fixed elevation (PPI scans, hence azimuthal technique), while assuming a two-dimensional cartesian geometric analysis plane (orthogonal to the beam pointing direction). The cartesian geometric analysis plane assumption is fulfilled for the ADL from the beginning, making the technique applicable. Differing from Frehlich et al. (2006), the technique is applied to the vertical wind in the work presented here, as in Lothon et al. (2006). According to Frehlich (2000), the analysis of the vertical wind is the most robust analysis scheme, since the mean vertical wind is zero and vertical wind velocity statistics are well behaved. For analysis, isotropy is assumed, therefore the length scales of the vertical wind in both transversal directions are assumed to be equal (e.g. in the horizontal x and y direction, see Lothon et al., 2006, for the differentiation between longitudinal and transversal length scales for the vertical wind). Consequently, no distinction is made in the PVA correction procedure for cross-wind vs. along-wind legs. The validity of this assumption is checked in Sec. 6.7. Based on the results by Frehlich and Cornman (2002), the averaging in the direction orthogonal to the beam is neglected and only the PVA in along-beam direction is considered (as is also done in Frehlich et al., 2006).

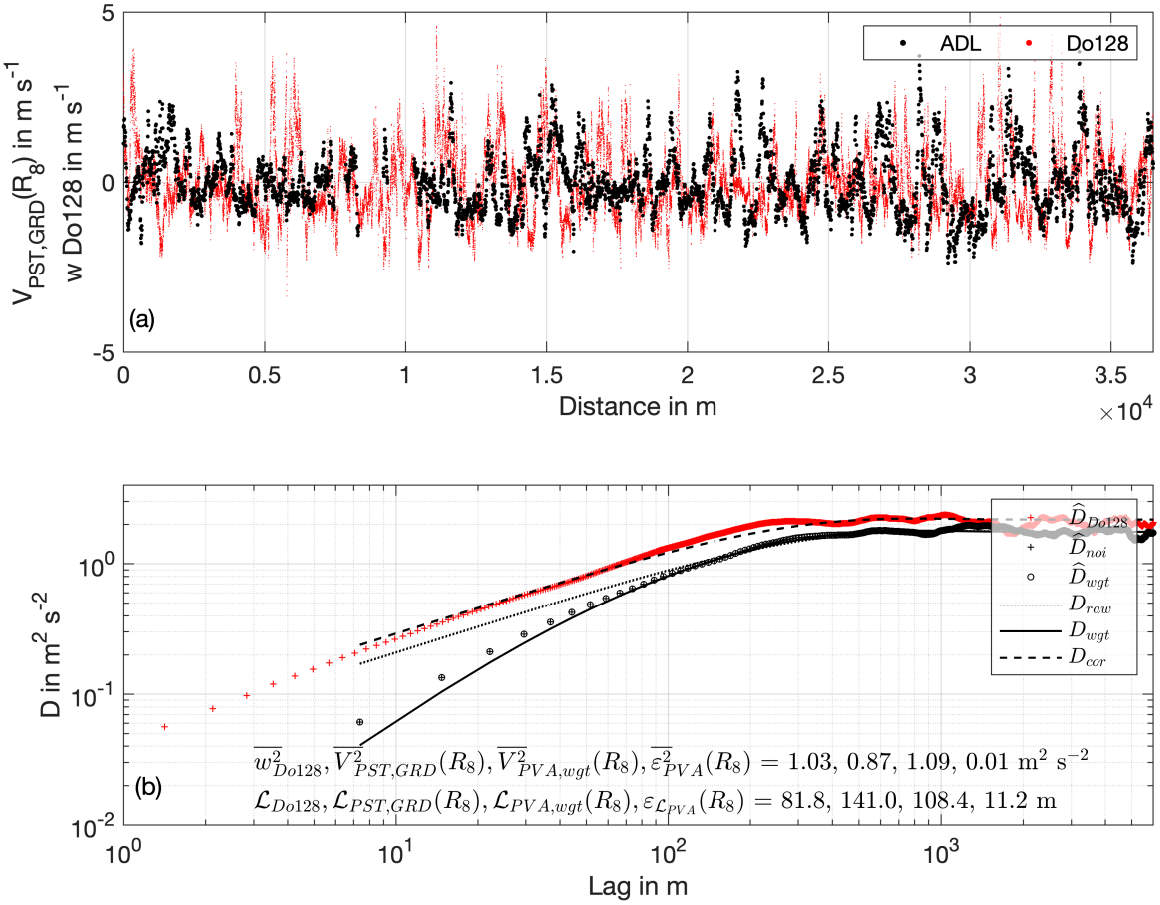

Fig. 6.18.: Vertical wind measurements and structure functions used for PVA correction for two flight legs during flight 4 discussed in Sec. 6.7. (a) Displayed is the motion and wind corrected ADL measurement for the range gate R_8 (10:46:30-10:55:15 UTC). Further displayed is the Do128 in-situ vertical wind measurement (10:58:30-11:07:45 UTC), conducted in close spatial and temporal proximity by approximately retracing the measurement path of R_8 at a similar vertical level. (b) Structure functions calculated using the vertical wind measurement and modeled using the modified azimuthal structure function technique based on Frehlich et al. (2006). Labels according to the specifications given in the text, measured structure functions are marked as points and their symbols denoted with a hat, whereas idealized structure functions are marked as lines. The Do128 structure functions extends to smaller separations due to the higher measurement frequency of 100 Hz compared to 10 Hz for the ADL.

In the following, the PVA correction procedure using the modified azimuthal structure function technique is illustrated and discussed for one representative example (Fig. 6.18). The PVA correction procedure is applied exemplary for the flight legs discussed in Sec. 6.7. For this flight section the path of the lowest lidar measurement level (R_8) was retraced by the aircraft after the initial measurement, in close spatial and temporal proximity. Thereby, additional in-situ data is available to illustrate and verify the PVA correction procedure, although the in-situ data cannot correspond to the lidar measured data in absolute values (details on the measurements are provided in Sec. 6.7). A maximum of 22 minutes passed between the two measurements (begin

of lidar measurement to end of aircraft measurement), a time period for which stationarity can usually assumed to be valid. Unfortunately, no co-located ground measurements were available to capture the temporal evolution of the wind at a fixed location due to the test character of the flights. Besides the temporal evolution, comparison of the absolute values has to be interpreted with care, as the measurements were conducted in the presence of vertical gradients, thereby small vertical displacements can cause differences. Nevertheless, basic visual analysis shows that the in-situ data of the aircraft exhibits more and stronger fluctuations at small scales compared to the lidar data (Fig. 6.18a). The aircraft measured in-situ data shows a higher vertical wind variance of $1.03\,\mathrm{m^2\,s^{-2}}$ compared to the PVA compromised value of $0.87\,\mathrm{m^2\,s^{-2}}$ for the lidar. Further, the aircraft measured integral length scale of 82 m is noticeably shorter than the lidar measured, PVA compromised integral length scale of 141 m. The same analysis scheme as discussed already is also shown in Fig. C.27 during the first leg (for the same time as the measurement of $\overline{V^2_{PST,GRD}}(R_8)$), but for the uppermost lidar range gate R_1 and the aircraft in-situ data measured 400 m above. Due to the temporal co-location of measurements, some agreement between the vertical wind time series is observed. However, the aircraft was flying close to the BL top (approx. $0.9z_i$, below the cloud base), therefore the vertical gradient in the variance and integral length scale profile makes a direct comparison impossible. Further, during the second part of the leg, the lidar measures noticeably different turbulent structures compared to the aircraft (the vertical wind magnitude exceeds the ADL measurement uncertainty by an order of magnitude and the structures are vertically persistent towards the ground).

The PVA effect is corrected based on idealized structure functions, which are generated utilizing a von-Karman turbulence model (see Appendix A.8 for the associated formulas). For these idealized structure functions, the PVA effect can be taken into account based on existing theory (Frehlich, 1997; Frehlich et al., 1998; Frehlich and Cornman, 2002; Frehlich et al., 2006). Based on the theory by Frehlich et al. (1998, 2006), idealized PVA-compromised structure functions D_{wgt} are generated utilizing the structure function of a von-Karman model D_{raw}. The idealized PVA-compromised structure function takes into account the pulse length and range gate length of the lidar setup (Sec. 3.1) as well as the measurement geometry (two-dimensional plane). Then, the real, observed structure function $\hat{D}_{wgt}$ (also PVA-compromised, noise corrected from $\hat{D}_{noi}$) is fitted to the idealized PVA-compromised structure function. For calculation of the structure function $V_{PST,GRD}$ is used. Figure 6.18b shows that the idealized PVA-compromised structure function generally matches with the observed structure function down to small separations. This general agreement gives confidence in the applicability of the von-Karman model (discussed in detail in the next section) as well as the simulated PVA effect using the azimuth structure function technique. Another example is given in Fig. C.27b, where the idealized com-

promised structure function and the lidar measured structure function differ more noticeable. The PVA correction works as expected (variance increase, integral length scale decrease), but due to the deviations the PVA correction should be treated with caution (also shown by the increased length scale uncertainty). The LSQ-fit contains $\overline{V^2}_{PVA,wgt}, \mathcal{L}_{PVA,wgt}$ as free parameters, which are fitted as a function of the structure function separation s. The LSQ-fit is bounded by the coefficient limits 0-10 $\mathrm{m^2\,s^{-2}}$ for the variance and 0-1500 m for the integral length scale (0-2000 m for the outer scale), structure function estimates in the range $10\,\mathrm{m} < s < 6000\,\mathrm{m}$ are used (a higher upper bound is used compared to ground-based studies due to the faster aircraft speed compared to the wind advection speed). Thereby, the best suited parameters to describe the observed measurements are determined from the model, but *with* consideration of the PVA effect. The PVA-corrected values of $\overline{V^2}_{PVA,wgt}, \mathcal{L}_{PVA,wgt}$ are also those which can then be used to reconstruct a not PVA compromised structure function D_{cor}, using the original von-Karman model again without accounting for the effect of PVA (for illustrative purpose only). This reconstruction is the structure function as it would be measured by an in-situ measurement not affected by PVA. Comparison with the aircraft in-situ measured structure function $\hat{D}_{Do128}$ (Fig. 6.18b) shows that the reconstructed structure function D_{cor} agrees better than the original structure function D_{raw}, both in level and in shape. Observable differences are related to temporal and/or spatial displacements between the measurements as well as uncertainty in the applied PVA correction procedure.

Figure 6.18b also clearly shows the advantage of using the azimuthal structure function technique compared to the longitudinal technique. Due to the calculation orthogonal to the beam pointing direction, the structure function is resolved down to very small separations of 7 m (compared to every 72 m for along-beam calculation). Besides the aforementioned higher vertical resolution of 72 m compared to several hundred meters, the critical, strongly varying region of the structure function is captured better and can be evaluated.

Applicability and uncertainty of the pulse volume averaging correction procedure

Before relying on the PVA correction procedure the general applicability of the von-Karman model in describing the turbulence present needs to be evaluated (besides investigating the LSQ-fit quality). Applicability is checked by comparing the $\overline{V^2}_{PVA,raw}, \mathcal{L}_{PVA,raw}$ values obtained from the raw transversal von-Karman model fit (without accounting for PVA) to those obtained from the direct calculation, as is done by Brugger (2014).

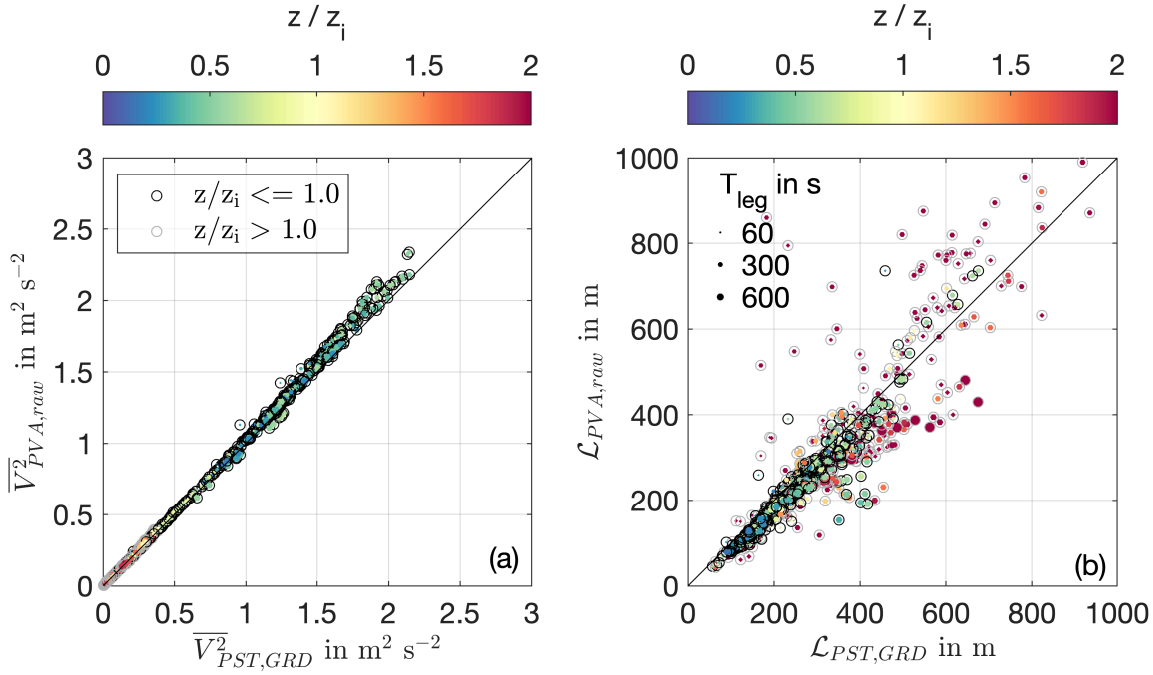

Fig. 6.19.: Comparison of (a) the variance and (b) the integral length scales obtained using the direct calculation and the fit of the observed structure function to the raw transversal von-Karman model. ADL measurements are displayed for all altitudes on all nadir legs with a minimum leg length of 60 s. The leg length is displayed through the circle area. The altitude of the individual measurements compared to the BL height is color-coded, the von-Karman model is only expected to yield reliable results inside the BL. Vertical deviations from the 1:1 line show the uncertainty $\overline{\varepsilon^2_{PVA}}$ associated with the PVA correction procedure, due to questionable applicability of the idealized von-Karman model. For points circled in gray the PVA correction cannot be applied due to their location above 1.0 z_i and the associated non-applicability of the von-Karman model.

Before getting started, a remark on the transversal versus longitudinal von-Karman model and its application in the Frehlich et al. (2006) procedure is necessary: One should not confuse the neglect of averaging in the direction orthogonal to the beam with the calculation of the measured structure function in a direction transversal to the vertical wind (orthogonal to the beam pointing direction, in the horizontal, therefore transversal to the vertical wind direction). Consequently, for the results presented here, the applicable model structure function is the transversal von-Karman model as in Frehlich et al. (2006) (which is derived from the longitudinal model but includes an additional term). However, as discussed in Appendix A.8, PVA correction of the vertical wind warrants a modification of the Frehlich et al. (2006) technique, as the length scales defined therein are in a longitudinal direction, whereas they are measured in a transversal direction when using the vertical wind. Hence, a factor 2 is necessary to capture the difference between the length scales (Houbolt et al., 1964).

The results obtained from the raw model fit are shown in Fig. 6.19 for the raw transversal von-Karman model using ADL measurements for all flights from all nadir legs and all levels with a minimum data availability of 60 s (including turns gives very similar results but is not displayed here to avoid overcrowding the figures). The variance shows very high agreement over the full range with very limited deviation from the 1:1 line, independent of the available leg length. This finding is promising, as it shows that the raw transversal von-Karman model is generally able to represent the measured variance well. Good agreement is also observed for the integral length scale, although noticeably more scatter is present. The scatter, as well as unreliable length scales $> 500\,\mathrm{m}$, are associated with measurements above the BL (the BL height is determined from the lidar measurements themselves and aircraft profiles in close proximity). In this region, the von-Karman model cannot yield reliable results as no homogeneous turbulence is present (therefore also the definition of a length scale itself becomes futile). As expected, the PVA correction can thereby not be applied outside the BL (and this is also unnecessary, as the non-PVA corrected lidar measurement are reliable in this region already, as is shown in Sec. 6.4).

It should be noted, that the agreement between the directly calculated quantities and the raw von-Karman model fit, shown here, is substantially better than the comparison shown in Brugger (2014). Most likely, the improvement is due to the more reliable statistical properties of the analyzed time series when using airborne measurement platforms, as well as the more appropriate usage of the transversal von-Karman model for the vertical wind, including the factor 2 mentioned before. Non-negligible differences between aircraft and ground-based measurements for vertical wind statistics are also documented by Adler et al. (2019). Stationarity is better fulfilled when using airborne platforms compared to ground-based measurements due to the shorter measurement periods. Similarly, the applied Taylor hypothesis is also better fulfilled, as the aircraft speed is considerably higher than the wind advection speed. As another side effect, the autocorrelation of the data is thereby also reduced, providing for more reliable statistics (as an averaging over more eddies is possible).

To my knowledge, no theory exists on the estimation of the uncertainty associated with the PVA correction procedure. For this reason, a simple approach is chosen here: The uncertainty $\overline{\varepsilon_{PVA}^2}$ associated with the PVA correction procedure, e.g. due to questionable applicability of the idealized von-Karman model, is approximated by the difference between the raw von-Karman model fit and the direct estimates.

All of the procedures outlined above were also implemented using the autocovariance functions instead of the structure functions, based on the theory provided by Frehlich (2000) and Brugger (2014). The results (not shown for the sake of brevity) are very similar to those obtained using the structure function based approach, giving confidence in the applied algorithms. As ex-

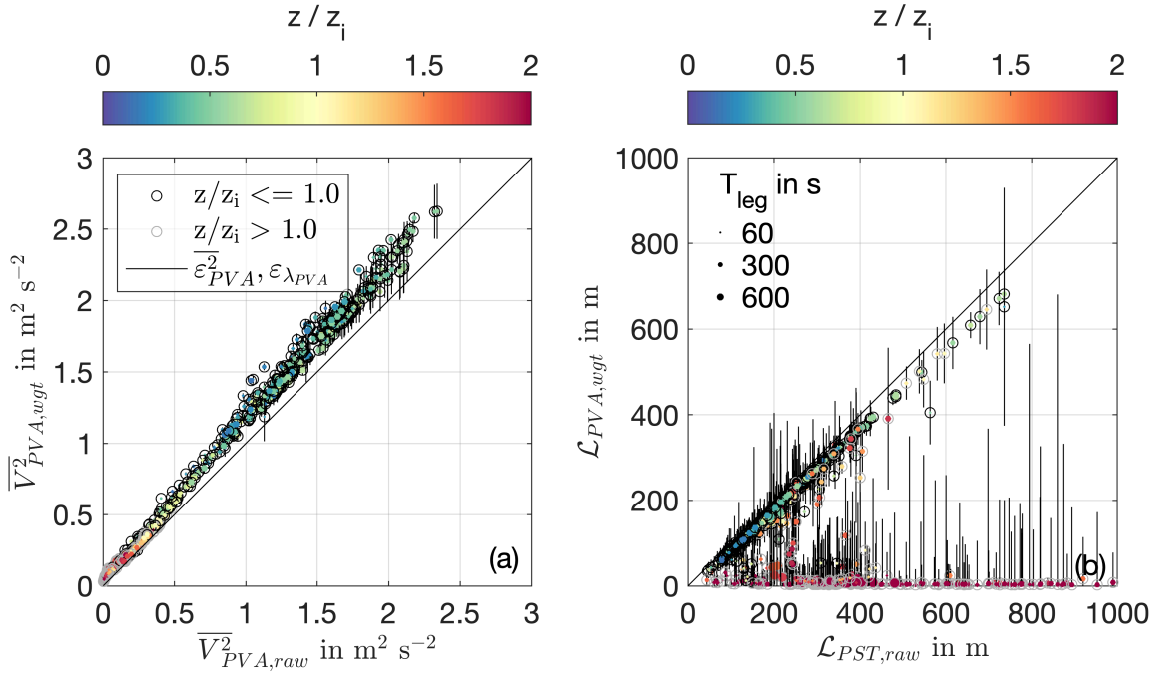

Fig. 6.20.: Effect of PVA correction procedure on (a) the variance and (b) the integral length scales. The PVA corrected values are shown in comparison to the raw von-Karman model estimates. ADL measurements are displayed for all altitudes on all nadir legs with a minimum leg length of 60 s. The leg length is displayed through circle area. The altitude of the individual measurements compared to the BL height is color-coded, the von-Karman model is only expected to yield reliable results inside the BL. Deviations from the 1:1 line show the PVA correction $\overline{\beta^2_{PVA}}$. For points circled in gray the PVA correction is not applied in the final data due to their location above $1.0 z_i$ and the associated non-applicability of the von-Karman model. Vertical bars illustrate the uncertainty $\overline{\varepsilon^2_{PVA}}$ associated with the PVA correction procedure based on the raw von-Karman model applicability.

plained by Brugger (2014), small existing differences in the obtained results between using the structure and autocovariance function can be explained due to the different weighing of points by the fit routines. The more stable results reported by Brugger (2014) when using the autocovariance based approach can not be confirmed here, possibly due to the already very stable results when using the structure function based approach.

Effect of pulse volume averaging correction procedure

After evaluation of the applicability of the von-Karman turbulence model the PVA correction can be applied to all measurements inside the BL. The results are shown in Fig. 6.20, comparing the values obtained from the raw von-Karman model fit to those using the PVA weighted model. The comparison to the directly estimated values is shown in Fig. C.28.

Both show results as expected and are very similar between each other. The PVA corrected variances are considerably increased on the order of $0-30\%$. The relative increase is stronger for lower variances, where the lidar 'loses' more of the variance due to PVA because these are associated with shorter integral length scales on average. Consequently, in absolute numbers, the PVA correction increase is almost constant over a wide range. For the variance, the PVA correction magnitude is well above the uncertainty associated with the application of the von-Karman model (which is $\overline{\varepsilon_{PVA}^2} = 0.024\,\mathrm{m^2\,s^{-2}}$ on average for nadir legs). The rather large corrections stress the importance of performing a reliable PVA correction. The PVA corrected integral scales are reduced by $0-20\%$ through the PVA correction procedure (average uncertainty 57 m). As explained before, the von-Karman model is not applicable outside the BL and the definition of an integral scale becomes futile, consequently the integral scales obtained from the PVA deviate in this region. The results show that the deviation already starts in the region of the BL top under some circumstances. Potentially, the deviation is due to more organized convection as well as entrainment and detrainment processes becoming important, which may invalidate the von-Karman model (Fig. C.27). This finding is an interesting subject for further study. Application of the PVA correction can be associated with larger uncertainties. However, uncertainties in the region $100-400\,\mathrm{m}$ (encountered most often) are small and decrease with decreasing integral scale, making the PVA correction also reliable.
The procedures developed here can be applied to a scanning system in the same manner without fundamental changes. However, for a scanning system there is potentially faster movement of the beam through the atmosphere which needs to be considered (depending on the scan pattern, e.g for wind profiling scans). Due to the faster movement, the condition $\Delta h << \Delta p$, formulated by Frehlich and Cornman (2002), may be violated for some scan pattern (nadir stares are still unproblematic and results are expected to be comparable). Therefore, fulfillment of the condition needs to be monitored for a scanning system. In case of violation, averaging orthogonal to the beam pointing direction (in the direction of the beam movement) also needs to be considered in the PVA correction. Consideration is possible using a similar framework as presented here.

6.2.VIII. Data processing and uncertainty due to finite measurement length

In order to obtain meaningful atmospheric variances based on the lidar data, a number of additional procedures need to be a applied to the data, besides the quality control procedures explained in Sec. 6.2.III. First, data gaps can be present in the lidar data for various reasons. Second, detrending and windowing of the time series should occur to mitigate undesired edge

effects in the spectral analysis. Third, systematic and random uncertainties of the statistical properties remain due to the finite length of the observed time series (substitution of the ensemble mean with the temporal mean). An overview on the basis of the time series and spectral analysis techniques discussed here is provided in Sec. 3.4 and therefore not repeated here.

Data gaps in the lidar data can be present for various reasons, e.g. too low SNR or other quality filtering reasons, blocking due to clouds and instrument calibration (e.g. closed shutter). Whereas the data gaps are less critical for wind profiling, they can prevent the calculation of meaningful statistics during nadir legs. Besides the data gaps, blocking due to clouds also introduces a systematic effect, as the region of the strongest updrafts are often missed due to their position below the clouds. Possible remedies and ways to deal with these data gaps are discussed by Kiemle et al. (2007, 2011) in-depth and the approach outlined by Kiemle et al. (2011) is followed here. Consequently, only cloud free legs and legs with data gaps smaller than the size of the dominant turbulent structures are used for statistical analysis during nadir legs (generally, very few data points are missing inside the BL if no blocking by clouds occurs). Any existing data gaps are zero-padded and the variance reduction thereby is compensated with a factor N/N_{valid}, both for the total variance as well as for the spectral decomposition. Thereby, N is the number of theoretically available data points and N_{valid} is the number of available data points. Nevertheless, the sudden jumps due to zero padding can result in elevated spectral noise levels (it should be noted that the data series cannot be pasted together as this would destroy the phase relationships). However, due to the generally high data availability in the BL and the few instances of occurrence, based on the results by Kiemle et al. (2011), this effect is deemed negligible compared to the other uncertainties discussed.

Before application of the spectral analysis the data needs to be conditioned carefully, in order to avoid undesired edge effects due to the applied technique but while preserving the atmospheric features. Generally, data conditioning includes straight-forward removal of the mean and linear trend also conducted in as a part of this work. More caution needs to be exerted when applying a data window to suppress spectral leakage, as too strong windowing can alter the statistical properties of the time series. Based on the recommendation by Stull (1988), a Tukey window is applied here which smoothes 10 % of the data edges with a tapered cosine function ($r = 0.2$). The weights of the Tukey window are defined through the following relation:

$$W(n) = \begin{cases} 0.5\{1 + cos(\frac{2\pi}{r}[n - r/2])\}, & 0 < n < \frac{r}{2} \\ 1, & \frac{r}{2} < n < 1 - \frac{r}{2} \\ 0.5\{1 + cos(\frac{2\pi}{r}[n - 1 + r/2])\}, & 1 - \frac{r}{2} < n < 1. \end{cases} \tag{6.5}$$

The Hanning window (a special case of the Tukey window with $r = 1$) applied by Kiemle et al. (2011) was found to alter the statistical properties too strongly for the time series investigated. As for the data gaps, a portion of the variance is lost due to application of the window function which needs to be restored again.

Overall, the following description to correct the variance correction (in the spectrally decomposed version based on Kiemle et al. (2011) is obtained:

$$S(f) = \frac{2}{\Delta f} \frac{N}{N_{valid}} \frac{N}{\sum_{k=1}^{N} W(k)^2} |S_{FFT}(W \cdot x, f)|^2. \tag{6.6}$$

The proper functionality of the spectral decomposition is controlled by comparing the variance computed from the spectral decomposition with that calculated directly from the processed time series (also after undergoing the data conditioning), which gives equal results as should be. It should be noted that in the data processing no filter function (e.g. high-pass filter or others) is applied to the time series. When investigating variances (compared to fluxes) the spectral gap does not always exist, nor is it necessarily clearly defined (Lenschow and Sun, 2007). Further, the contribution to variances from the meso-scale can be interesting to analyze and should not be omitted.

Even after careful data conditioning, systematic and random uncertainties remain because the observed time series is finite. These uncertainties are calculated based on Lenschow and Kristensen (1984) and Lenschow et al. (1994) as explained in Appendix A.9. For the variance, the systematic underestimation $\overline{\beta_{sys}^2}$ is added to the measured variance. However, to my knowledge no uncertainty estimation for the estimation of $\overline{\beta_{sys}^2}$ is available from literature, therefore the random uncertainty $\overline{\varepsilon_{sys}^2}$ is equated to the systematic underestimation. Besides the systematic uncertainty, the much larger random uncertainties are also calculated based on Lenschow et al. (1994), both $\overline{\varepsilon_{ran}^1}$ for the mean and $\overline{\varepsilon_{ran}^2}$ for the variance. The random contribution only decreases with $\sqrt{\frac{\tau}{T}}$ compared to $\frac{\tau}{T}$ for the systematic contribution, which makes it the more prominent source of uncertainty for reasonably long flight legs.

Results for the variance are shown in Fig. C.29 for all altitudes on all nadir legs with a minimum data availability of 60 s. Turn legs show a similar behavior, although with slightly elevated uncertainty levels, because on average their leg length is shorter compared to nadir legs. As expected, the systematic offset is much smaller ($< 0.1\,\mathrm{m^2\,s^{-2}}$) compared to the random uncertainty, which can be considerable up to $0.45\,\mathrm{m^2\,s^{-2}}$. Both measures are functions of the absolute variance observed and thereby increase with increasing variance. At any given variance, longer leg lengths result in more accurate measurements. Nevertheless, some scatter exists between equal leg lengths due to the varying integral length scales present. The average random uncertainty

of the variance is $0.095\,\mathrm{m^2\,s^{-2}}$ for nadir legs, although of course in practice individual estimates are available and used for all range gates on all legs. The average random uncertainty of the mean is $0.11\,\mathrm{m\,s^{-1}}$ for nadir legs (not shown).
The data processing methods for random and systematic uncertainty estimation applied here can be used in the same manner for a scanning system. Finally, all relevant sources of error for ADL measurements inside the BL have thereby been assessed and quantified through their associated uncertainties. Following this, the retrieval accuracy for wind profiling and nadir measurements is evaluated in a comparative analysis and by a comparison to aircraft in-situ measurements.

6.3. Wind profiling measurement summary

A synopsis of the ADL measurement accuracy is provided in Table 6.3 for the first order moment. The wind profiling accuracy is dependent on the radial velocity measurement accuracy during turns, which is used as the input for wind profiling (as the aerosol movement (V) is negligible). The error contributions (I-IV) matter for the radial velocity accuracy, from these the combined errors of (I)+(II)+(IV) can be conservatively bounded and further refined based on the measured ground return velocities (Sec. 6.2.IV).
Overall analysis of the results obtained and discussed in the previous sections demonstrates that none of the error sources degrades wind profiling accuracy considerably. The largest potential contributor of systematic offsets and random uncertainty is the beam pointing-angle. However, this potential uncertainty contribution is identified as a strong overestimation. Most of the uncertainty in (II) is due to the insufficiently explored b_y-direction, which does not influence the motion correction process noticeably due to the small aircraft velocities in this direction. This reasoning is validated through the measured ground return velocities, which provide a conservative bound of the actual error present and show small deviations. The second largest potential influence is the INS accuracy. If no ground return is available for refined correction (e.g. wind profiling above clouds) the magnitude of uncertainty introduced by the INS can be up to $0.12\,\mathrm{m\,s^{-1}}$, which is still low. If the refined motion correction procedure is applied, using the residual ground return velocities for correction of INS error, the uncertainties due to the terms (I-IV) become negligible.

Tab. 6.3.: Overview of typical maximum (<) or average (≈) systematic offsets and random uncertainties associated with mean wind measurements by the ADL system investigated in this work. Uncertainty contributions (I-IV) determine the radial velocity measurement accuracy, which can be controlled using the measured ground return velocities. Uncertainty contributions (V-VIII) are relevant for retrieval of the atmospheric vertical wind from nadir observations. All values present conservative estimates, the refined values are obtained by applying the procedures outlined in Sec. 6.2.

Source of error	Systematic offset $\overline{\beta^1}$ in $m^1\,s^{-1}$				Random uncertainty $\overline{\epsilon^1}$ in $m^1\,s^{-1}$				Main factors, comments	Indiv. leg, range gate availability
	Raw		Refined		Raw		Refined			
	Nadir	Turn	Nadir	Turn	Nadir	Turn	Nadir	Turn		
(I) Time synchronization	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	Flight dependent	Yes, -
(II) Beam alignment	< 0.05	< 0.05	≈ 0	≈ 0	< 0.17	< 0.17	≈ 0	≈ 0	Installation dependent	Yes, -
(III) Radial velocity noise	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	@ 10 Hz measurement freq., SNR dependent	No, Yes
(IV) Motion correction	≈ 0	≈ 0	≈ 0	≈ 0	< 0.06	< 0.12	≈ 0	≈ 0	INS accuracy & ground availability dependent	Yes, -
Measured (I)+(II)+(IV)	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0.02	≈ 0.05	≈ 0	≈ 0	Error bound based on ground return	Yes, -
(V) Terminal fall velocity	< 0.01	< 0.01	< 0.01	< 0.01	≈ 0	≈ 0	≈ 0	≈ 0		No, No
(VI) Horizontal wind contribution	≈ 0.03	≈ 0.03	≈ 0.03	≈ 0.03	≈ 0.22	≈ 0.37	≈ 0.13	≈ 0.18	Wind profile & flight stability dependent	No, No
(VII) Pulse volume averaging	-	-	-	-	-	-	-	-	Not applicable for mean	-, -
(VIII) Finite time series	-	-	-	-	≈ 0.10	≈ 0.11	≈ 0.10	≈ 0.11	Leg length dependent, turbulence dependent	Yes, Yes

This evaluation is confirmed by an assessment of the mean vertical wind retrieval accuracy. The horizontal wind correction error, including the contributions (I-V), can be bounded conservatively through the deviation of the measured mean vertical wind from a $0\,\mathrm{m\,s^{-1}}$ mean vertical wind expected for continuity reasons, when measured over extended legs. Using the INS-based motion correction one obtains a bias of $0.03\,\mathrm{m\,s^{-1}}$ and a random error below $0.22\,\mathrm{m\,s^{-1}}$ for ADL measurements of the mean vertical wind. This accuracy can be improved by using the refined motion correction procedure, which leads to a slightly lowered bias and a substantially reduced random error of $0.13\,\mathrm{m\,s^{-1}}$. Together with the random uncertainty of $0.10\,\mathrm{m\,s^{-1}}$ due to analysis of a finite time series (VIII), it can be assessed that the ADL measurements of the mean vertical wind provide excellent accuracy under a wide range of conditions.

To capture the effect of the errors (I-IV) on wind profiling their associated uncertainties are propagated through the retrieval uncertainty estimation framework, in the same manner as is done for the radial velocity variance due to flow inhomogeneity. Besides the accuracy of the radial velocity measurements, the validity of the flow homogeneity assumption used for wind profile retrieval can present a major driver of AVAD retrieval error. This error was investigated in depth using the ADLS (Chap. 5), and the data-driven uncertainty estimation method developed there is applied to the real-world measurements here.

If the ADL retrieved wind profiles are compared to those from another measurement system (e.g. aircraft in-situ data), additional sources of disagreement can occur. These are measurements errors in the system used for comparison as well as differences due to spatial and/or temporal displacement of the compared wind profiles. In this work, the ADL retrieved wind profiles are compared to those obtained from the Do128 in-situ measurement system. For this purpose, aircraft vertical soundings were flown prior to or after ADL wind profiling turns for a number of profiles. During flight 1, the aircraft soundings were conducted by circling in proximity to the ADL wind profiles. During the other flights, comparisons were also made by using straight climbing or descending legs (at the cost of increasing distance between ADL and aircraft profile). Both flight types entail non-negligible in-situ wind measurement errors. Firstly, during vertical maneuvers the aircraft INS shows a systematic bias which influences the vertical wind measurement (see Sec. 6.2.IV). Second, during turns the in-situ wind measurement was found to exhibit periodic error in dependence of the heading (partially due to INS error but with further systematic contributions, potentially to an error in the nose-boom orientation). Both sources of error can be of considerable magnitude up to $1\,\mathrm{m\,s^{-1}}$. Besides the pure measurement system errors, differences between the ADL retrieved and Do128 in-situ measured winds can also be introduced by spatial and/or temporal displacements between the measurements.

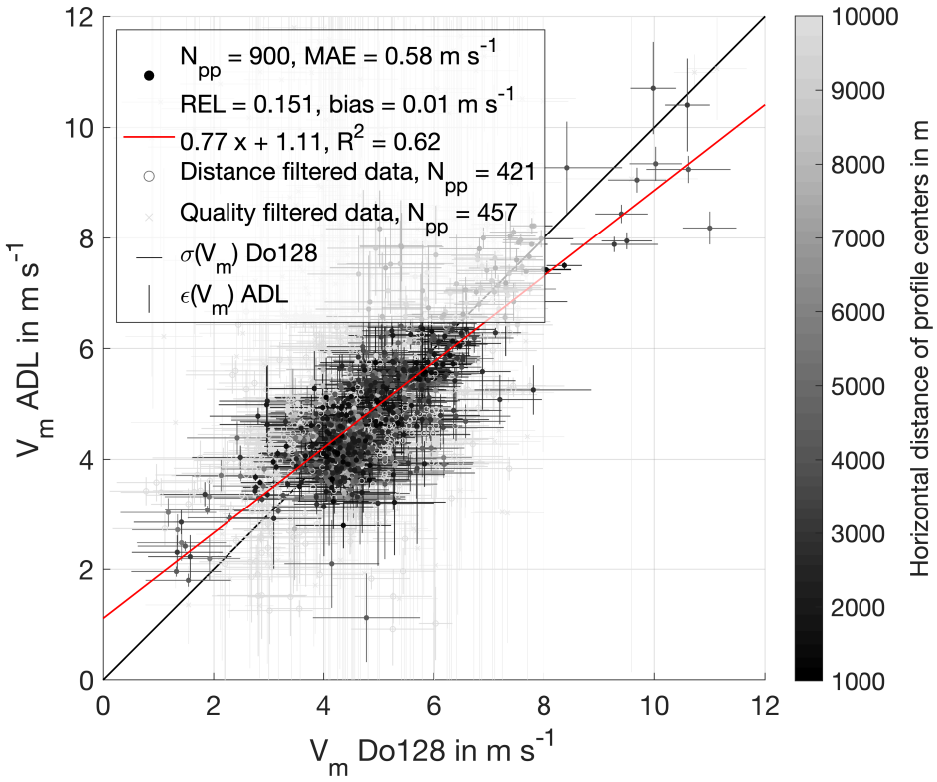

Fig. 6.21.: Comparison of all co-located wind profile points on all flights, showing Do128 in-situ measured and ADL remotely sensed wind speed from AVAD retrieval. For comparison, the aircraft measured in-situ wind profile is averaged to the lidar altitude resolution. The horizontal distance between aircraft and ADL profile is color coded, as agreement can only be expected in close horizontal proximity. Only profiles with a horizontal center distance <10 km are used for the calculation of statistical parameters. Only *CN*-based quality control is applied to the ADL retrieved wind profiles ($CN < 12$), to increase spread of the *CN* the ADL profiles are artificially sub-sampled for full 360° turns. Vertical bars illustrate the uncertainty estimate of the ADL retrieved wind speed based on the theory developed using the ADLS. Horizontal bars show the standard deviation of the Do128 in-situ measured wind speed during passage of the corresponding ADL height interval.

In summary, the additional sources of error and the differences introduced thereby can be much larger than the typical expected ADL wind profiling error, making an assessment of ADL wind profiling accuracy based on real-world comparisons difficult (hence the ADLS was developed). Nevertheless, the comparison is shown here for completeness and as some interesting features in similarity to the ADLS results can be detected, further, another very conservative bound of ADL wind profiling error can be obtained. No attempt is made to quantify and separate the above error contributions as a part of this work, as these sources of error can be neglected in the ADLS, where the data-driven uncertainty estimation applied here is developed and validated.

In some instances the aircraft flew more than one full 360° circle for wind profiling, because a directional change was linked to the profiling or for measurement consistency evaluation. For

comparison, all circles with more than one full circle are subdivided in order to retrieve multiple wind profiles. While these profiles are not statistically independent (because the wind field is very similar between circles), the subdivision and shorter rest-circles serve to create artificial, challenging measurement conditions by increasing the spread in the condition number. As a consequence, the measurement quality of the raw profiles is degraded and the applicability and reliability of the quality control criteria can be thoroughly investigated.

The comparison between Do128 in-situ measured and ADL retrieved vertical profiles of the horizontal wind is shown in Fig. 6.21 for the wind speed, in Fig. 6.22 for the wind direction and in Fig. C.30 for the u, v, w components. For comparison, the in-situ measured wind profile is averaged to the lidar altitude resolution. Therein, the measurement error present in the Do128 in-situ measurement, as well as the potential effect of spatio-temporal separation of the profiles, is approximated through a simple measure: the horizontal bars show the variability of the wind speed as measured by the Do128 during the passage of the respective lidar altitude bin. The vertical bars represent the uncertainty as given by the data-driven uncertainty estimation method developed in the ADLS when applied for the ADL. Further, color-coding reveals the horizontal separation of wind profiles used for comparison. For statistical evaluation only profile points with a separation distance <10 km are compared.

Despite the many additional sources of error, which are not attributable to the ADL measurement accuracy, a generally good agreement is observed. $N_{pp} = 900$ wind profile points are available for comparison with an MAE of $0.58\,\mathrm{m\,s^{-1}}$ (REL = 0.15, bias = $0.01\,\mathrm{m\,s^{-1}}$, $R^2 = 0.62$). A good agreement is also observed for the wind direction with an MAE of 7.4° (REL = 0.04, bias = 1.2°, $R^2 = 0.97$). The lower REL and higher R^2 are due to the much better explored range of wind directions compared to the wind speed where values cluster around $5\,\mathrm{m\,s^{-1}}$. The horizontal wind components show a similar MAE as for the wind speed and a higher $R^2 > 0.9$, due to the more spread distribution (the REL shows similarly high values as for the wind speed, but only due to the presence of values around $0\,\mathrm{m\,s^{-1}}$). A meaningful comparison of the mean vertical wind retrieved from AVAD scans is not possible using the current comparison strategy, due to the biased Do128 vertical wind measurements. Nevertheless, small data-driven ADL uncertainty estimates suggest that some retrieved vertical winds could actually be meaningful. If all flown circles are used in full, instead of the subdivision after turning 360° (to artificially create challenging conditions and generate higher CN spread), the achieved quality metrics are very similar although with fewer points available for comparison (not shown; $N_{pp} = 613$, MAE = $0.60\,\mathrm{m\,s^{-1}}$, REL = 0.156, bias = $0.05\,\mathrm{m\,s^{-1}}$, $R^2 = 0.63$). This observation proves two things: First, adding more circles does not necessarily increase measurement accuracy. This is expected, because the turbulence and thereby violation of the flow homogeneity assumption is

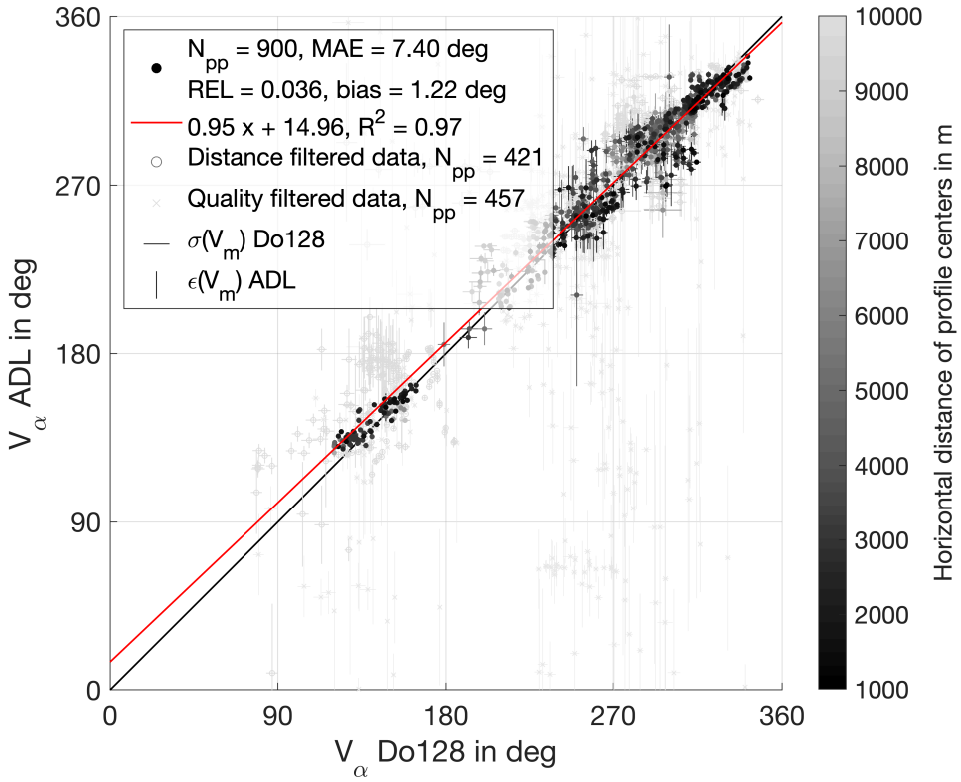

Fig. 6.22.: Comparison of in-situ measured Do128 and ADL remotely sensed wind direction for all co-located profiles on all flights. Displayed quantities correspond to Fig. 6.21 except that the wind direction is shown.

very similar between two circles. Because this violation is the major driver of wind profiling uncertainty, the error between two subsequent profiles at the same location is very similar and not decreased. Second, quality control with the *CN* works reliably. Despite the more challenging conditions, due to the artificial *CN* degradation, a similar level of wind profiling accuracy is achieved.

Another deeper look into the quality control, especially at low wind speeds, provides further interesting insight (Fig. C.31). Many of the profiles sample low wind speeds $< 5\,\mathrm{m\,s^{-1}}$. Using the ADLS, this low wind speed region was identified as potentially problematic due to the erroneous mapping of vertical wind into horizontal wind for inhomogeneous flow conditions. Associated with the problematic wind speed retrieval at low wind speeds, a number of interesting features in close agreement with the ADLS results can be observed. The quality controlled data (and the LSQ-fit) shows a positively biased ADL retrieval at very low wind speeds. This bias is caused by the erroneous mapping of vertical wind into horizontal wind (Sec. 5.1.2). Unfortunately, the optimized system setup and retrieval strategy suggested from the ADLS results (Sec. 5.3) cannot be applied for the prototype version of the ADL because no scanning capabilities are available. Additional circles do not mitigate the problem, since no independent new

information is obtained (as discussed above already). Consequently, here the intended scanning ADL can provide considerable added value.

As in the ADLS, the biased retrieval only occurs for the wind speed, whereas the retrieved u,v,w components remain unbiased. As already observed in the ADLS results, a clear dependence of the retrieval bias on the correlation coefficient exists (Fig. C.31a, b, c). Consequently, quality control with a stricter R^2 threshold increases the agreement (smaller MAE), but also increases the bias at lowest wind speeds due to stronger in-phase filtering for the vertical wind, as expected. Besides, the number of retrievable profile points is strongly reduced. For example, setting a $R^2 > 0.9$ threshold, the obtained quality parameters are $N_{pp} = 578$, MAE = $0.48\,\mathrm{m\,s^{-1}}$, REL = 0.11, bias = $0.10\,\mathrm{m\,s^{-1}}$, $R^2 = 0.76$.

Hence, the added value of the data-driven uncertainty estimation becomes clear. It allows for assessment of the reliability of the ADL retrieved values without the application of R^2-based quality control, thereby preserving a large number of retrievable points and avoiding a potentially increased bias. Using the data-driven uncertainty estimation method, the actual added value of the ADL data is also much higher than suggested by the MAE (besides the strong overestimation of the MAE due to the above discussed additional error contributions). For example, during the morning flight on 18 June, ADL retrieved wind profiles are generally more reliable due to the little turbulence present. Thereby, at this time observed differences between aircraft and ADL wind profile points are traceable to Do128 wind measurement error or insufficient co-location (see also Fig. C.26). Different conditions were encountered during the afternoon flight on 18 June, when low wind speeds combined with inhomogeneous flow were present. Thereby the uncertainty of ADL retrieved profiles is increased, which is captured by the ADLS developed data driven uncertainty estimation method, see Fig. C.32 for an example.

Tab. 6.4.: Overview of typical maximum (<) or average (≈) systematic offsets and random uncertainties associated with the measurement of atmospheric variances using the ADL system investigated in this study. All values present conservative estimates, the refined values are obtained by applying the procedures outlined in Sec. 6.2. For the correctable systematic offsets (III,VII,VIII) the random uncertainty quantifies the uncertainty associated with the correction technique. Please note that for all relevant uncertainties an individual leg-based estimation procedure is available. Thereby, the expected systematic offsets and random uncertainties can be quantified for all measurements, given the specific measurement conditions present. Consequently the leg-based uncertainty estimates can be substantially lower than the typical maximum and average values shown here.

Source of error	Systematic offset $\overline{\beta^2}$ in $m^2\,s^{-2}$				Random uncertainty $\overline{\epsilon^2}$ in $m^2\,s^{-2}$				Main factors, comments	Indiv. leg, range gate availability
	Raw		Refined		Raw		Refined			
	Nadir	Turn	Nadir	Turn	Nadir	Turn	Nadir	Turn		
(I) Time synchronization	< 0.005	< 0.005	≈ 0.001	≈ 0.004	-	-	-	-	Flight dependent	Yes, -
(II) Beam alignment	< 0.033	< 0.37	≈ 0.001	≈ 0.004	-	-	-	-	Installation dependent	Yes, -
(III) Radial velocity noise	0-0.2	0-0.2	≈ 0	≈ 0	-	-	< 0.01	< 0.01	@ 10 Hz measurement freq., SNR dependent	Yes, Yes
(IV) Motion correction	< 0.006	< 0.085	< 0.002	< 0.015	-	-	-	-	INS accuracy & ground availability dependent	Yes, -
Measured (I)+(II)+(IV)	≈ 0.003	≈ 0.017	≈ 0.001	≈ 0.004	-	-	-	-	Error bound based on ground return	Yes, -
(V) Terminal fall velocity	< 0.01	< 0.01	< 0.01	< 0.01	-	-	-	-		No, No
(VI) Horizontal wind contribution	≈ 0.016	≈ 2.657	≈ 0.003	≈ 0.058	≈ 0.066	≈ 5.701	≈ 0.062	≈ 0.153	Wind profile & flight stability dependent	Yes, Yes
(VII) Pulse volume averaging	0-0.5	0-0.5	≈ 0	≈ 0	-	-	≈ 0.024	≈ 0.019	Turbulence dependent	Yes, Yes
(VIII) Finite time series	0-0.1	0-0.5	≈ 0	≈ 0	≈ 0.095	≈ 0.105	≈ 0.095 + 0.016	≈ 0.105 + 0.024	Leg length dependent, turbulence dependent	Yes, Yes

6.4. Vertical wind variance measurement summary

ADL measurements of the vertical wind variance are subject to a number of systematic offsets, some of them correctable, as well as random uncertainties discussed in detail in Sec. 6.2.I-6.2.VIII. A synopsis of the magnitude of the individual terms which demand consideration is provided in Table 6.4. To my knowledge, this table and the preceding analysis present the most complete investigation of ADL vertical wind variance measurements up to date. Previously, no studies existed on the possibility of ADL variance measurements from flights inside the BL.
From the error sources (I-IV), the beam pointing-angle (II) maps into the biggest potential systematic uncertainty contribution to the raw radial velocity variance, especially during turn maneuvers. Fortunately, an analysis of the measured ground return velocity variance identifies the beam pointing uncertainty as a strong overestimation (it is constructed as a worst-case scenario). Further, the ground return velocity analysis allows for a conservative estimation of the measurement error bounds associated with the sum of the terms (I)+(II)+(IV). Results show that the systematic variance contribution by these terms is small, it is further reduced when the refined motion correction procedure is available. The uncorrelated noise (III) presents an unwanted additional variance contribution to any lidar measurement and can be of non-negligible magnitude. Differing from the terms (I), (II) and (IV), its systematic effect can be corrected in the refined analysis with little remaining random uncertainty (using the first-lag autocorrelation estimation method detailed in Sec. 6.2.III).
A number of additional error sources (V-VIII) demand consideration, if the vertical wind variance is to be extracted from the measured radial velocity variance. Of these, the uncertainties associated with the terms (VI-VIII) can be of considerable magnitude, whereas the terminal fall velocity (V) of the scattering aerosol is typically negligible. The horizontal wind contamination due to non-nadir beam pointing (VI) is of non-negligible magnitude and it's effect cannot be corrected. Nevertheless, both the systematic offset and random uncertainty introduced by it can be conservatively estimated using the framework provided by Strauss et al. (2015). In a refined version, the magnitude of the uncertainty is reduced in the present work through removal of the mean horizontal wind contribution, which also enables variance measurements during turns. PVA (VII) leads to a non-negligible systematic decrease of the ADL measured variance compared to that obtained by an in-situ measurement. For the refined measurements, the systematic PVA effect is considered and corrected using the modified azimuthal structure function technique based on Frehlich et al. (2006). The random uncertainty associated with the PVA correction procedure is quantified through the appropriateness of the raw von-Karman model fit, which yields acceptable accuracy. Last, the measured variance is subject to systematic and

random uncertainties due to the finite length of the time series (which are not ADL specific but also present in in-situ measurements, e.g. those of the Do128). These effects are evaluated based on the Lenschow et al. (1994) framework. Especially the random uncertainty can be of considerable magnitude and demands careful consideration, its magnitude can be reduced through using longer averaging periods.

Overall, the conducted analysis shows that ADL measurements of the vertical wind variance are subject to a number of non-negligible systematic offsets and random uncertainties. Nevertheless, application of a number of refined, measurement condition dependent and range gate specific, uncertainty estimation methods allow for reliable variance measurements.

For nadir analysis of the vertical wind variance, an in-situ comparison for accuracy assessment is even more difficult than for wind profiling. The preferable option would be a co-located measurement at the point of the lidar measurement, e.g. through a second aircraft flying below. Unfortunately, such measurements are expensive and difficult to execute, they were not available during the prototype test flights analyzed here. Therefore, another approach is chosen, in which the lidar path was retraced by the Do128 in close spatial proximity after completion of an initial leg. Results from this ADL-aircraft comparison are analyzed in detail in Sec. 6.2.VII for the PVA correction and Sec. 6.7 for the general comparability. Unfortunately, retracing the lidar path in such a way still suffers from temporal and co-location differences. Additionally, the low flight level of the aircraft does not allow for lidar measurements and valuable flight time is used, which is then not available for other measurement purposes. Therefore, only one such instance was conducted during the test flights.

Here, another possibility is explored in which the ADL measurements from an upper range gate are compared to the in-situ measurements of the aircraft. Thereby, measurements are conducted at the same time, but > 400 m apart vertically. Obviously, this comparison is problematic due to the different altitudes at which the measurements are conducted. For flights inside the BL, the aircraft often flew close to the BL top, consequently the lidar measured in the middle and lower part of the BL. In this case, for typical variance profiles as in Sec. 6.5, the ADL should measure systematically higher values of vertical wind variance. The exact amount of variance difference between ADL and aircraft depends on the vertical profile and cannot be generalized easily.

The results of the Do128-ADL comparison are shown in Fig. 6.23. As expected, the ADL measured systematically higher variances when the aircraft is flying close to the BL top. The systematically higher values are well above the uncertainty present in the ADL and in-situ measurements. The systematic difference decreases for legs on which the aircraft is flying more towards the middle of the BL. If the aircraft flies only slightly above the BL ($< 1.5 z_i$), the ADL

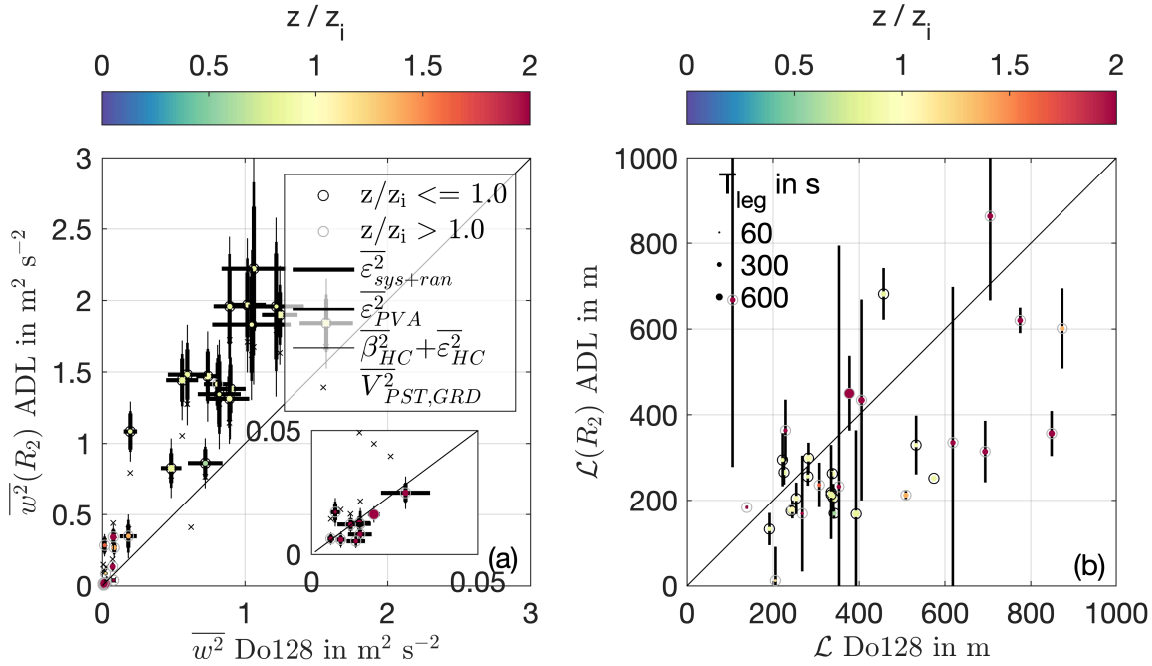

Fig. 6.23.: Comparison of Do128 in-situ measured and ADL remotely sensed (a) vertical wind variance and (b) integral length scale for all nadir legs on all flights. The ADL measurements are shown for the second range gate, which is located 470 m below the aircraft. The leg length is displayed through the marker size, longer legs yield more reliable estimates. The height of the aircraft flight leg in relation to the BL height is color coded. Black crosses show the raw variance values before noise and PVA correction. Noise correction is applied to all ADL measurements. PVA correction is only applied for ADL measurements inside the BL, as the von-Karman model is not applicable outside (Sec. 6.2.VII). Please note that the systematic offset between aircraft and ADL measurement is because the aircraft often flew close to the BL top. Thereby, the ADL measurements sample systematically higher variances compared to the aircraft due to the vertical variance gradient inside the BL (see Sec. 6.5).

still measures inside the BL or en-/detrainment zone, it thereby also shows elevated levels of variance compared to the in-situ measurements.

Nevertheless, an interesting opportunity opens for flight legs during which both the aircraft and the ADL measure well outside the BL. In these cases, both the aircraft and the ADL should measure comparable low levels of variance. Indeed, this agreement can be observed in the very low variance region $< 0.05\,\mathrm{m^2\,s^{-2}}$, which is challenging to measure for both systems. The good agreement illustrates the capability of the ADL to measure variances over a wide range of levels. Further, it stresses the importance and quality of the noise correction procedure. The raw, non-noise corrected ADL measurements show systematically higher values of variance. After noise correction, these fall closely within the region of the 1:1 line. The relatively long range gate setting of the WTX lidar performs better at low turbulence levels compared to other lidar settings with shorter range gates due to the different scaling of the random velocity error for

low and high turbulence (Frehlich, 1997; Frehlich and Cornman, 2002). Last, this comparison also shows that the in-applicability of the von-Karman model outside the BL does not impede measurement quality there.
The integral length scale comparison shows more scatter and elevated levels of ADL uncertainty. This is especially true outside the BL, where the definition of the integral scale becomes problematic. No method to estimate the uncertainty of the aircraft measured integral length scale is known to the author, therefore it cannot be easily quantified. Despite the overall more scatter, a weak correlation can nevertheless be observed, which is more promising than the results of Brugger (2014). Further, the method of quantifying the ADL length scale uncertainty post PVA correction through the raw von-Karman model deviation appears promising, as larger deviations are mostly associated with larger uncertainties.
In summary, the previous sections have shown that reliable wind profiling and variance measurements are possible using the investigated ADL system. Based on these results, three exemplary case studies are analyzed from the conducted test flights in the following. Hence, the new insights and added value which can be provided by a high-resolution ADL system on a medium-range turboprop aircraft are demonstrated for the first time.

6.5. Application 1 - The structure of turbulence in a BL topped with cumulus humilis above flat terrain

In this section, the ability of the prototype ADL system to provide high resolution vertical wind measurements throughout the full atmospheric BL is used for a demonstrative retrieval of vertical wind statistics in a cloud-topped BL. For this purpose, flight leg 3 of flight 1 is analyzed, which was conducted above flat terrain across a cloud-free section of a cloud-topped BL, with shallow cumulus humilis present elsewhere (see Sec. 6.1 and Fig. C.33). Figure 6.24 provides an overview of the obtained raw lidar data as well as the motion corrected and post-processed vertical wind retrieval. The lidar signal quality is sufficient to enable vertical wind retrieval throughout the full measured column. Additionally, an undisturbed ground retrieval is available throughout the leg, allowing for application of the refined motion correction procedure. Consequently, after horizontal wind post-processing, the lidar is able to resolve the structure of BL turbulence in great detail and a number of features are visible with the naked eye already.

Boundary layer structure

The BL is clearly identifiable through its associated turbulence and the increased SNR (Fig. 6.24c, d). In contrast, the free troposphere above shows reduced SNR (although far from being

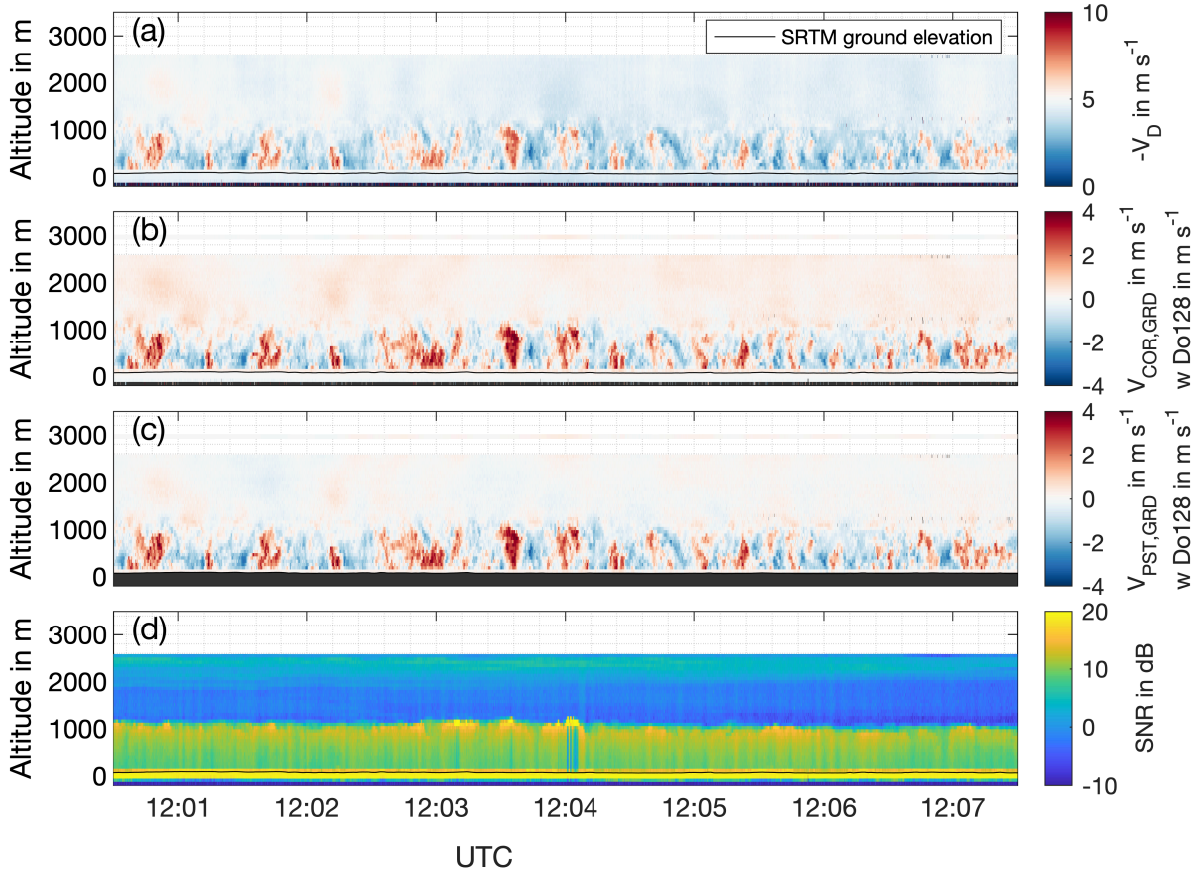

Fig. 6.24.: Illustration of ADL measured data in a BL topped with shallow cumulus humilis for leg 3 during flight 1 on 10 July 2019. Displayed are: (a) Uncorrected lidar radial velocity measurement. (b) Refined motion corrected lidar radial velocity measurement (color coded curtain). Additionally, the aircraft vertical wind measurement is extended vertically by ±50 m and displayed at flight level as a color coded line. (c) Refined motion and horizontal wind corrected lidar measurement (color coded curtain). Aircraft vertical wind measurement as in panel above. (d) Lidar SNR. The ground elevation determined using a DEM is displayed as a black line in all panels.

unusable) and no turbulence. Based on SNR analysis, the BL height is fairly homogeneous with $z_i = 1100\,\mathrm{m} \pm 100\,\mathrm{m}$, although a few updrafts appear to lift it slightly. The BL turbulence is organized into distinct updrafts of varying shapes and spacing. Some up- and downdrafts extend all the way from the surface to the BL top, whereas others do not (they might elsewhere, but not within the 2D-plane which was overflown). The area covered by the updrafts is smaller than those covered by the downdrafts, consequently, the vertical wind in the updrafts is stronger than in the downdrafts. At 12:04 UTC a small cloud filament was overflown, however the filament was so thin that the lidar was able to penetrate. The filament is not associated with a very distinct updraft, possibly pointing to a cloud in the process of decay. Shortly before, however, a very distinct updraft was passed with vertical winds up to $+4\,\mathrm{m\,s^{-1}}$. It appears that the updraft

has not quite reached the BL top yet, potentially leading to the formation of a clouds shortly after passage. Already, an increase in SNR is apparent at the top of the updraft, potentially due to the onset of condensation.
Another interesting feature is visible in the free troposphere. There are defined, extended regions of weak up- or downward motion of air, which extend from the lidar measurements to the in-situ measurements of the Do128. Consequently, these are not measurement artifacts but rather physical motions. Analysis of the in-situ measured vertical wind and temperature series reveals a 90° phase shift of the time series (not shown). Thereby, these features are identified as weak gravity waves which are induced by the BL turbulence and propagate upward. However, they are not directly linked to the underlying up- and downdrafts at first glimpse.

Vertical profiles

The vertical wind measurement quality on the analyzed transect is sufficient to allow for the calculation of the statistical moments and the associated uncertainties using the data processing and correction scheme outlined in Sec. 6.2. The obtained results are shown in Fig. 6.25. Fig. 6.25a shows the previously discussed data in a different display version, which illustrates the nearly constant measurement height of the range gates throughout the leg. In this case, already the second range gate above the ground level allows for vertical wind retrieval.
Fig. 6.25b shows that the mean vertical wind over the analyzed leg is close to $0\,\mathrm{m\,s^{-1}}$, within the limits of the measurement uncertainty. The integral length scale profile shows largely similar integral length scales $< 200\,\mathrm{m}$ throughout the BL. As expected, a decrease is observed towards the ground. Towards the top of the BL, the raw length scale estimates tend towards larger values, whereas the PVA corrected values tend towards zero. As discussed in Sec. A.8 the definition of a length scales becomes problematic outside the BL, therefore these values should be treated with caution.
Fig. 6.25c shows the associated second order moment profile of the vertical wind. The shape of the profile closely resembles that analyzed by Lenschow et al. (1980) for a baroclinic, convective BL. The profile shows that the maximum variance is reached in the lower half of the BL, where the full eddy size is reached and organization of updrafts into plumes occurs, at considerable magnitude $\overline{w^2} > 2\,\mathrm{m^2\,s^{-2}}$. Towards the ground, the profile shows a decrease in variance, as the size of the largest eddies which contribute to variance production becomes limited by the ground proximity. Towards the BL top the variance level also decreases, as on the investigated transect no convective clouds are present above, which could further strengthen the updrafts through latent heat release (see Sec. 6.7 for such an example). On the contrary, entrainment of

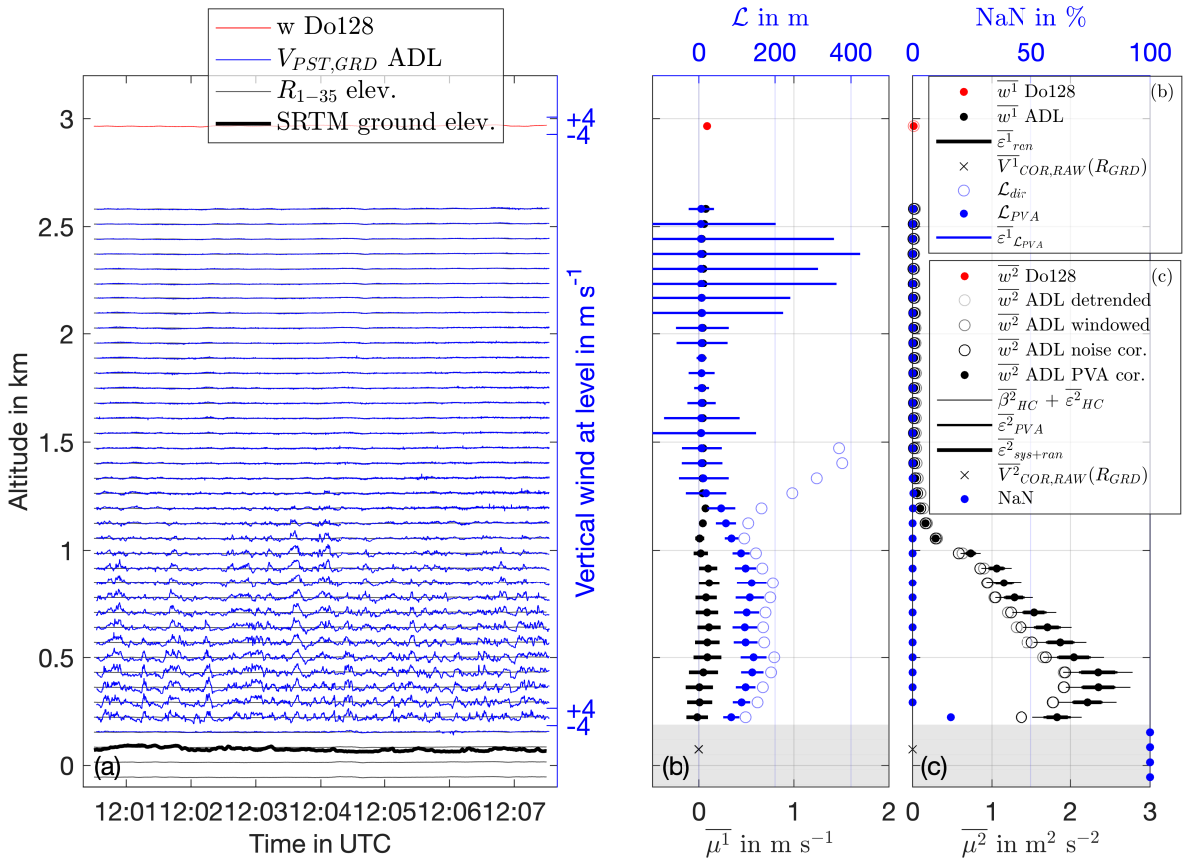

Fig. 6.25.: ADL measured vertical wind and associated profiles for transect shown in Fig. 6.24. (a) Time-height cross-section of ADL and Do128 measured vertical wind. (b) Corresponding profile of the mean vertical wind and integral length scale as measured by the ADL and Do128. (c) Corresponding profile of the vertical wind variance and fraction of unavailable data. All symbols and lines according to the respective legends, with additional information provided in the text.

air from the free troposphere with low associated turbulence levels and higher stability leads to a decrease of the profile already well below z_i. Once z_i is reached, the stable layering associated with the inversion above the BL suppresses turbulence, leading to a further rapid drop in the variance profile towards $0\,\mathrm{m^2\,s^{-2}}$. In this region, the PVA correction is not performed anymore due to the in-applicability of the von-Karman model (Sec. 6.2.VII) and the removal of the lidar uncorrelated noise becomes important for reliable variance retrieval (Sec. 6.2.III).

All relevant drivers of systematic offsets and random uncertainty in the variance retrieval are displayed, none of them reaches a magnitude which prevents a meaningful retrieval. Random and non-nadir beam pointing induced uncertainties are largest and approximately of equal magnitude with up to 20% of the measured variance. The PVA correction corrects the systematic

lidar underestimation offset and increases the variance by approximately 20%, once again highlighting the need for a reliable correction.

Spectra

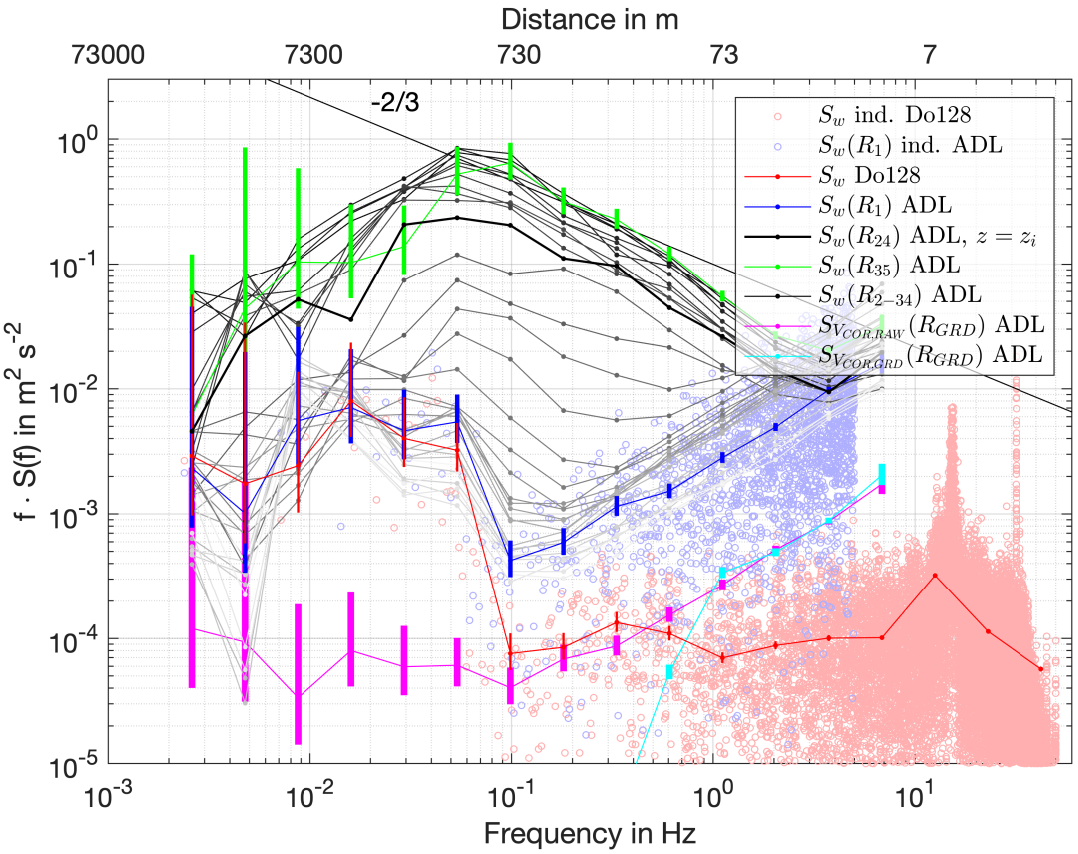

Fig. 6.26.: Spectral decomposition of the vertical wind time series shown in Fig. 6.25. The ADL measured spectra are binned logarithmically equidistant and faded with increasing height from the ground. The ADL range gate closest to the ground is colored in green. The ADL range gate in the region of the BL height is marked bold. The first ADL range gate is colored blue and the underlying raw spectral data is shown. The Do128 in-situ measured spectrum is shown in red, also together with the underlying raw spectral data. The magenta line shows the spectrum of the measured ground return velocity after raw motion correction. The cyan line shows the spectrum of the ground return velocity obtained after using the measured ground return for refined motion correction. The vertical lines represent the 90% confidence intervals of the averaged spectral estimates. They are only shown for the specific spectra mentioned to avoid overcrowding and as they are similar for the other ADL spectra. The confidence intervals are calculated assuming a chi-squared distribution and a Daniell spectral estimator based on Storch and Zwiers (1999).

The measured time series and associated variance can be analyzed further using spectral decomposition, which provides additional interesting insight (Fig. 6.26). While the aircraft measurement allows for calculation of only one spectrum at flight altitude, the ADL measurements

provide continuous spectra down towards the ground. ADL spectra from neighboring range gates are similar to each other both in level and shape, especially inside the BL. However, the spectral curves should not be over-interpreted, because the measurements are conducted in close vertical proximity with strongly correlated turbulence. Thereby, because the measurements are not independent, the statistical uncertainty in the spectral shape and level is not reduced by conducting multiple measurements at similar altitudes. Nevertheless, some distinct features are identifiable and traceable to physical characteristics of the observed turbulence structure of the vertical wind.

Inside the BL a distinct spectral peak, corresponding the the largest eddy size and turbulence production, is observed due to the $f \cdot S(f)$ display. The location of the peak shifts from smaller wavelengths at the ground towards longer wavelengths at the top of the BL. As stated above, the individual curves exhibit a similar shape and level due to the vertically correlated turbulence. The level of the spectra decreases towards the BL top (linked to the decreasing variance profile, see Fig. 6.25). In the region associated with wavelengths > 350 m the ADL spectra inside the BL show an approximately $-2/3$ decay, which is expected in the inertial subrange for the chosen display. For smaller lengths the ADL spectra drop off more rapidly due to the effect of PVA, which results in a loss of detected variance at the small scales by the lidar. Hence, the importance of the PVA correction procedure is demonstrated once again (the overall level of variance is PVA corrected, but the spectral shape is not). A weak uptick is observed in the lidar spectral level at the smallest lengths (highest frequencies), caused by the presence of uncorrelated noise in the lidar measurement. As the SNR decreases above the BL, the uncorrelated noise level increases. This increase is also detectable in the spectra, as the spectral lines above the BL exhibit a +1 slope characteristic for uncorrelated noise (Mayor et al., 1997). If a direct $log[S(f)]$ display style is chosen instead of the $log[f \cdot S(f)]$ shown here, the variance level in this region remains constant and can also be used for uncorrelated noise estimation (see Sec. 3.4, Frehlich and Cornman, 2002).

Above the BL, the lidar spectral levels are again similar at much smaller absolute values. A spectral peak is consistently observed at much longer wavelengths, potentially due to the presence of gravity waves discussed above. Interestingly, further strengthening confidence in the ADL measurements as well as the gravity wave hypothesis, the spectra of uppermost ADL range gate and Do128 in-situ measurement agree very well in this region, both in shape and level. As the in-situ measurement does not suffer from uncorrelated noise, no increase in spectral level is observed towards higher measurement frequencies. Nevertheless, some artifacts are present, e.g. a nose-boom oscillation frequency at approximately 10.5 Hz.

Additionally shown is the spectrum of the measured ground return velocity, both for the raw and refined motion correction. No concerning spectral features are observable for the measured raw ground return and the overall variance level is very low. Thereby, the high quality of the raw motion correction procedure in straight and level nadir flight is proven yet again and the gravity wave hypothesis is hardened further. As mandated by the correction, the spectral levels of the refined motion correction drop to even smaller levels of variance for frequencies lower than 1 Hz. Above 1 Hz raw and refined variance levels are similar due to the 1 s moving average applied for refined motion correction (Sec. 6.2.IV). The increase in spectral variance levels at higher frequencies, due to uncorrelated noise, is also clearly identifiable, albeit at much lower levels compared to the atmospheric spectra due to the higher SNR.

6.6. Application 2 - Spatially resolved wind observations in moderately complex terrain

In this section, the potential of the prototype ADL system to provide spatially resolved wind measurements of both the horizontal and vertical wind in moderately complex terrain is demonstrated. To this end, results from flight 2 and 3 are analyzed above the Harz region, capturing both the temporal and spatial development of the BL as well as its link to orographic features. The horizontal wind field encountered during the flights was complex and evolving throughout the day, related to the passage of a convergence line which was situated over central Germany in the afternoon. The synoptic situation is discussed in Sec. 6.1, a snapshot of the horizontal wind field from a NWP model is provided in Fig. C.23.

Figures 6.27 and 6.28 show the vertical wind across the Harz region, as measured on multiple transects along a similar ground path at different points in time. The associated SNR is shown in Fig. C.34 and C.35. Figure C.36 provides a 3D illustration of one such transect and the orography in the region. While some transects explore the full mountain range as well as the surrounding areas, others are shorter and focus on the main mountainous region due to flight time considerations or, in the case of the afternoon above cloud transect, visual flight rule limitations not permitting cloud penetration.

The obtained lidar data quality is high, with strong return signal also being available in the free troposphere. Thereby, a number of interesting features are captured which are discussed in the following to highlight the added value of an ADL.

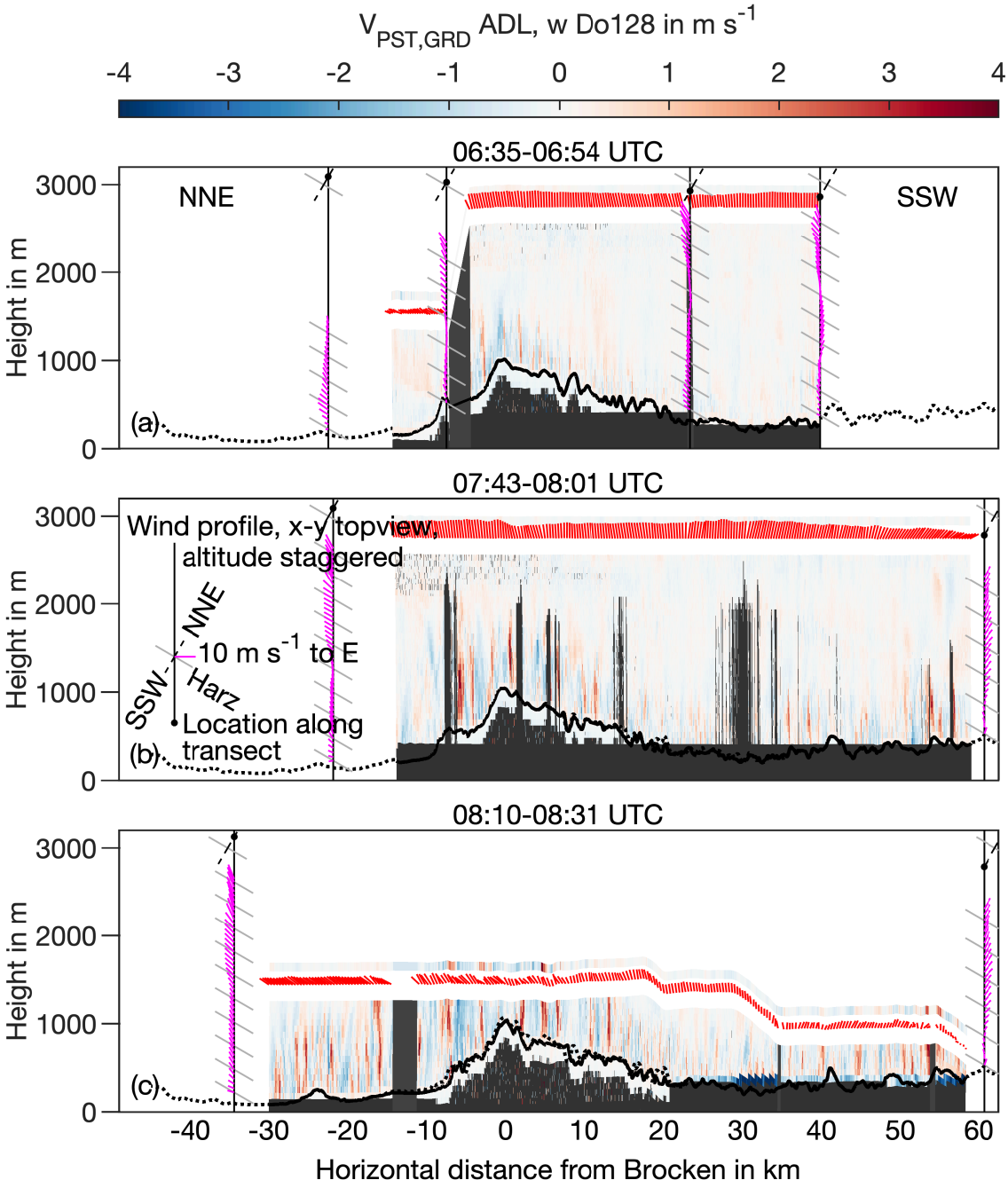

Fig. 6.27.: ADL retrieved and Do128 in-situ vertical wind for three nadir transects across the Harz mountain range during the morning flight on 18 July. The position of the transect is illustrated by the dotted line in Fig. C.16. The cross-sections are centered with respect to the highest peak called 'Brocken' and were conducted at different times during the flight, displayed above each panel. Data gaps are due to non-nadir maneuvers being flown or clouds/ground blocking the lidar signal. Two ground heights are displayed, as the ideal, intended cross-section of the Harz was not always exactly maintained by the aircraft. As usual, the solid line shows the SRTM ground elevation at the position of the lidar range gate intercepting the ground. Additionally, the dashed line always shows the ground elevation along the 009.95 °E, 51.0 °N – 011.18 °E, 52.5 °N transect for reference. Additionally, the quivers show the horizontal wind speed and direction obtained from ADL wind profiling and the direction of the Harz mountain range using an x-y topview. The quivers are staggered according to the profile altitude. In the same manner, the Do128 in-situ measured horizontal wind is shown along the flight path. A scale for the wind quivers is provided in panel (b).

Morning situation

The first transect was conducted early in the morning, before the generation of convective clouds above the mountain range (Fig. 6.27a). Cloud formation of shallow cumulus humilis occurred approximately 10 minutes after the crossing, as observed by the flight scientist. ADL retrieved data shows that the horizontal wind in the surroundings of the Harz region is blowing with 2 to 10 $\mathrm{m\,s^{-1}}$, predominantly from south-easterly directions. An exception is the altitude region from 1 to 2 km, where wind speeds are even lower and the direction is more southerly. The wind profile variations align with structures observed in the SNR, pointing to the presence of multiple layers with different flow regimes in the residual layer (Fig. C.34a). At lower altitudes, the wind direction is more ridge-parallel, whereas at upper altitudes a stronger southerly wind component introduces some cross-ridge component. The horizontal wind at flight level is mostly homogeneous along the flight path from southerly directions. The observed BL structure aligns with theoretical expectations. Moderately strong updrafts exist above the mountain range with vertical winds $< 3\,\mathrm{m\,s^{-1}}$ reaching up to 600 m AGL. In the surrounding, less complex terrain, BL development has not advanced as far yet and the BL height is still shallow < 200 m AGL. Above the highest elevations, the updrafts reach higher, are stronger and spaced further apart compared to the weaker, shallower and closer updrafts above the surrounding area. Little to no alignment of the updrafts with the individual topographic peaks is observed, likely due to the mostly along-ridge horizontal wind at lower altitudes.

A similar transect was flown again approximately one hour later (Fig. 6.27b). By then, the BL evolution has clearly progressed, consistent with theoretical expectations (De Wekker and Kossmann, 2015). The horizontal wind is still comparable to the previous transect. Interestingly, a weakening of the horizontal wind speed is observed in the in-situ measurements when crossing the Harz ridge-line. Unfortunately, ADL wind profiles are missing to evaluate whether this is a persistent effect at all altitude levels. Here, a scanning system would be beneficial as it allows for continuous wind profile retrieval along the flight path. The flight scientist noted the development of a few cumulus mediocris over the highest parts of the mountain range. Further, convective clouds also started to develop over some surrounding areas. The clouds are readily identifiable in the ADL measured data as they block the signal. The high level clouds at a distance of 30 km from the Brocken are not connected to the BL, instead they are shallow altocumulus. The BL height is identified to be approximately 1000 m AGL above the Brocken and approximately 500 m AGL in the surrounding areas (Fig. C.34b). Updrafts have strengthened compared to the previous transect and now exceed $+4\,\mathrm{m\,s^{-1}}$ in a few places. A field of convec-

tive clouds not connected to the Harz mountain range, but with distinct updrafts, is overflown at a distance > 50 km from the Brocken.
Interestingly, in many cases the updrafts appear to be connected to weak vertical motions in the free troposphere above. As in the previous section, many of the weak vertical winds are consistently observed by both ADL and Do128, pointing to the generation of gravity waves by BL turbulence again.
After the second high altitude transect the flight track was approximately retraced at lower altitude, with the aircraft flying just below cloud base whenever BL clouds were present (Fig. 6.27c). As expectable, the measured data looks similar to the previous transect but without data gaps due to clouds. At flight altitude, the horizontal wind is weaker and changes to south-westerly directions at distances > 30 km from the Harz. Above the Harz, wind speeds are 5 – 8 $m\,s^{-1}$ from southerly directions, inducing some across-ridge flow at lower altitudes. Vertical winds exceed +4 $m\,s^{-1}$ in some places and the strongest updrafts below clouds are now captured. Although a direct comparison of measured updrafts is hardly possible, due to the temporal evolution as well as spatial mismatch, similar features are detectable. Some updraft penetrate all the way to the flight level and are visible in the in-situ data, whereas other do not (yet). The altocumulus field at a distance of 30 km from the Brocken does not show any noticeable updraft structures. In the vicinity of the Brocken, downdrafts have strengthened and extend over a larger domain compared to the previous transect. Hiwever, it is impossible to tell whether this strengthening is due to advection, temporal evolution of the horizontal and/or vertical winds or slight spatial mismatch between the flight tracks. Here, additional ground-based measurements as well as usage of the autopilot for most accurate flight track retracing is recommended for future campaigns.

Afternoon situation

A fully different situation was encountered during the afternoon flight (Fig. 6.28), when widespread cumulus fields were present. In most parts, the cumulus fields consisted of cumulus mediocris with a vertical extent of approximately 1000 m, while in some parts (connected to an upper level shading by cirrus clouds) only cumulus humilis with a smaller vertical extent were present. During the flight, stronger westerly winds developed across the Harz and altered the horizontal wind field situation, especially on the northern side of the ridge. The modeled horizontal wind field is shown Fig. C.23, the temporal development was connected to the advancing convergence line mentioned above.

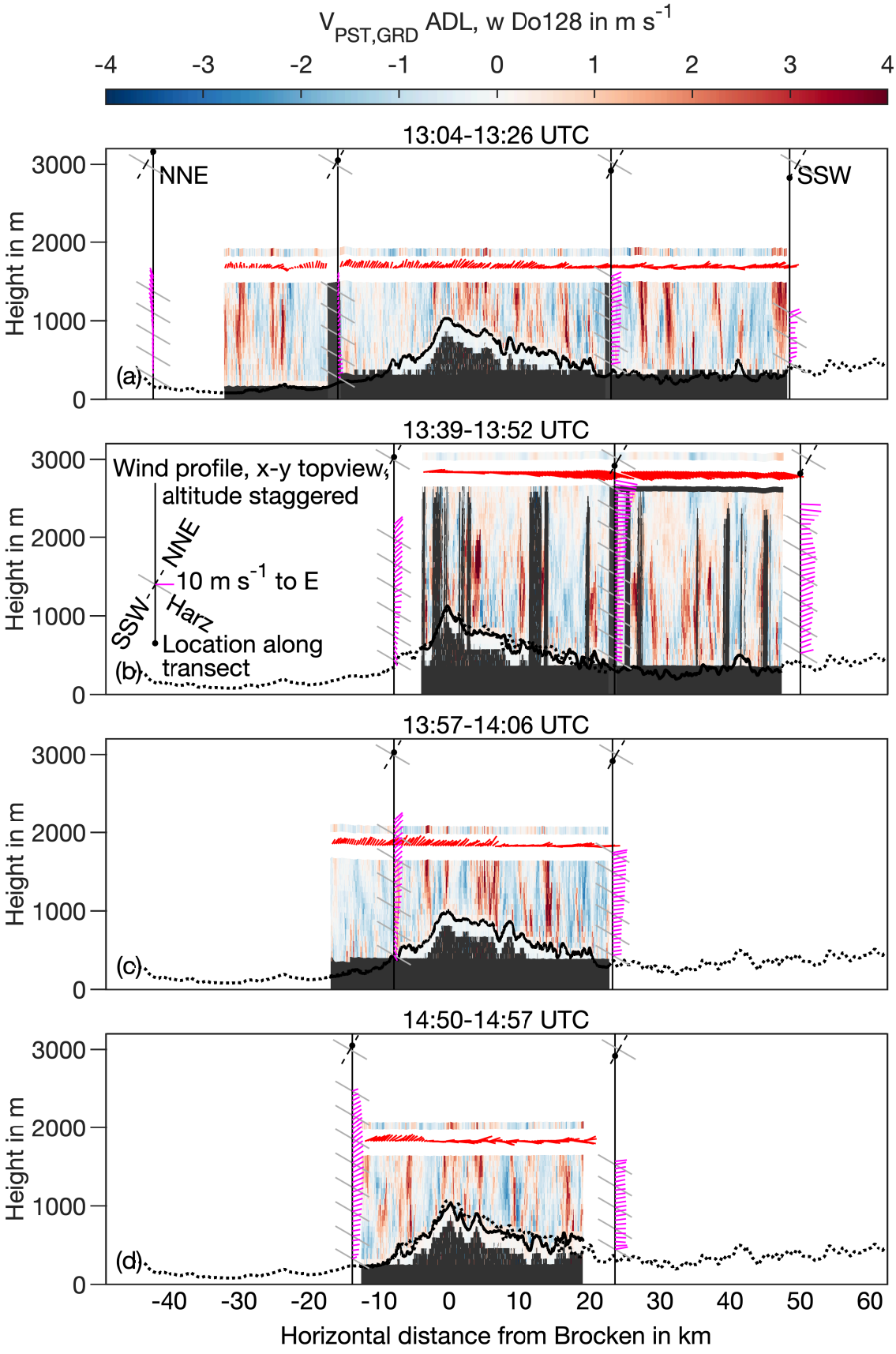

Fig. 6.28.: Same as Fig. 6.27 but for four nadir transects across the Harz mountain range during the afternoon flight on 18 July. The position of the transect is illustrated by the dotted line in Fig. C.17.

A first transect was conducted below clouds base across the same path as previously (Fig. 6.28a). A noticeable difference in the horizontal wind on both sides of the Harz ridge is observed. On the northern side of the Harz, wind speeds are low at all altitudes and the wind direction varies (it is problematic to define in low wind speed situations in any case). Above the Harz mountain range, horizontal wind speeds are increased and blow from a south-westerly direction (strong cross-ridge component). Towards the south, at all altitudes the wind direction is westerly (more ridge-parallel, although still with a slight cross component), at wind speeds around $8\,\mathrm{m\,s^{-1}}$.

Compared to the morning transects the evolved BL structure is readily identifiable. Much stronger, wider and deeper updrafts dominate the scene. As before, some updrafts are linked to clouds above the flight path and therefore also present in the in-situ vertical wind measurements. Differing from before, the Do128 did not ascend with the BL top across the Harz, therefore the BL top was estimated to be 200-300 m above the aircraft by the flight scientist above the Brocken. Clearly, different turbulence structures are encountered along the flight path. Towards the north, a mixed regime with some high reaching updrafts but also a much calmer region is present. Above the main Harz mountain ridge, a much more 'fuzzy' scene exists, with horizontally more narrow but nevertheless strong updrafts, spaced closer together. Potentially, this is connected to the stronger horizontal across-ridge flow in this region and/or a superimposed mountain venting circulation (De Wekker and Kossmann, 2015). The updrafts are not connected to the ridges at first glimpse, although a more sophisticated 3D analysis is required to evaluate this further. Towards the south, the widest, strongest and most distinct updrafts are observed. The vertical wind exceeds $+4\,\mathrm{m\,s^{-1}}$ in large areas. Interestingly, whereas updrafts appear to exhibit a pronounces difference in structure and intensity across the region, this does not appear to be the case for downdrafts in the same manner. Downdrafts mostly cover larger spatial areas but are less intense, with vertical winds not exceeding $-4\,\mathrm{m\,s^{-1}}$, and usually being in the range $0-2\,\mathrm{m\,s^{-1}}$.

After obtaining a vertical in-situ sounding, the transect was retraced above the clouds (Fig. 6.28b). Unfortunately, the transect had to be terminated prematurely due to clouds exceeding the 3000 m flight level and descent was initiated after passage of the Brocken. Clouds tops regularly reach to 2500 m but exceeded 3000 m in other places as stated above. A consistent BL height is difficult to define and exhibits large vertical variation especially in the region of downdrafts. Nevertheless, only a slight influence of the Harz mountain range on BL height is identifiable, with an increase of approx. 500 m (Fig. C.35). This observation is consistent with previous observations (Kalthoff et al., 1998) and theoretical expectations (De Wekker and Kossmann, 2015). Horizontal winds increase to speeds $> 10\,\mathrm{m\,s^{-1}}$ above the BL and change to

a north-westerly direction (along ridge). A similar counter-clockwise veer of the wind direction above the Harz is observed when flying northward and wind speeds are lower on the northern side. The obtained vertical wind data matches well with the previously measured transect. Due to the blocking by clouds, the most vigorous updrafts are systematically missed. Despite this, a number of strong updrafts can be observed between the clouds which have not reached the BL top yet and therefore not induced cloud formation. The many observed updrafts structures not connected to clouds point to a rapidly evolving cloud scene. From this first analysis conducted here, no distinct subsiding shells can be observed around the cumulus clouds, although some are linked to downdrafts in the surroundings (similar observation are made during the more cloudy flight on 19 July but not discussed here for the sake of brevity).

Once again, the cumulus clouds are connected to wave-like vertical wind structures in the free troposphere above. During the afternoon flight wave formed cirrus clouds were observed above parts of the measurement domain at higher altitudes (Fig. C.37). Indeed, the phase shift of the in-situ temperature and vertical wind measurements is close to 90° again, confirming this hypothesis.

A third transect was subsequently conducted below cloud base, slightly higher than the first transect, due to the elevated cloud base at the start point (approx. 1 h after the first transect, 20 min after the above cloud transect). Wind speeds are slightly increased compared to the first transect, especially on the norther side of the mountain range. There is once again a distinct region of wind veer (at approx. 10 km distance from the Brocken), which has progressed further north compared to the first transect. The encountered turbulence scene is comparable to that during the first transect (Fig. 6.28c), with a more 'fuzzy' turbulent situation above the Harz. The areas in the direct vicinity of the Harz show extended areas of subsidence and more narrow and shallow updrafts, as is also seen on the first transect. Thereby, more evidence for a potentially observed, superimposed mountain venting process is obtained (De Wekker and Kossmann, 2015). However, the data base is still insufficient to evaluate whether this subsidence is a systematic effect or coincidental.

A fourth and last transect was conducted again 1 h later towards the end of the flight (Fig. 6.28d). The horizontal winds are now blowing from the southwest to west at speeds of about $5-8\ \mathrm{m\,s^{-1}}$ throughout the full mountain range (e.g. a slight cross-range component). Once again, a noticeable veer region is identifiable which has progressed further north compared to the first and third transect (at approx. 5 km distance from the Brocken). A similar scene as before is observed although the updrafts have weakened slightly. During this transect, a strong updraft surrounded by strong downdrafts on both sides is observed, although not connected to a cloud (at 10 km distance from the Brocken).

Synopsis

During the morning scene the different BL structure is readily explainable and matches well with the expected earlier onset of convection above mountainous regions compared to the surrounding lowlands (De Wekker and Kossmann, 2015). While no connection of the observed thermals to orography could be made here, more frequent horizontal wind observations (e.g. using a scanning system) and a more sophisticated three-dimensional analysis might allow to do so.

The afternoon scene is more complex, both in terms of the horizontal as well as the vertical wind. Wind speeds were generally higher than during the morning flight. Throughout the flight, a region of higher wind speeds from more westerly directions was observed on its passage over the Harz mountains. The observed vertical wind scenes are heterogeneous with different extent and spacing of updrafts observed between the Harz and its surroundings. Potentially, the differences are associated with the changing horizontal wind field. The differences could also be connected to an influence of the underlying topography, which is more complex in the Harz region compared to the surroundings. Thereby, a superimposed mountain venting circulation can be induced, which can explain the repeatedly observed, extended regions of subsidence to both sides of the mountain range. Further, the more shallow BL above the Harz region in absolute terms may also contribute. The BL height is only slightly influenced by the terrain height in the afternoon, as observed before (Kalthoff et al., 1998). Consequently, the absolute vertical extent of the BL actually decreases above the mountain ridge, as it does not fully compensate for the increase in the terrain height. Thereby, updraft structure and spacing can also be altered, as was seen in the morning. Further study and more extensive measurements are required to evaluate these hypothesis further and distinguish their effects. To do so, the linkage of airborne measurements with comprehensive high resolution ground-based measurements (Damiani et al., 2008; Kalthoff et al., 2013, besides others) is mandatory. The potential of a scanning ADL system appears promising for further research, as it can provide continuous horizontal wind observations along the flight track, which is not possible using the prototype ADL.

Similar to the previous section, the ADL measurements also allow for analysis of the statistical parameters associated with the turbulence present above the mountain range and its surroundings. However, due to the moderately complex terrain as well as the inhomogeneous meteorological conditions present, this is a complex undertaking and requires careful in-depth evaluation of the individual measurements. It is therefore not conducted here for the sake of brevity but remains an interesting option for future studies.

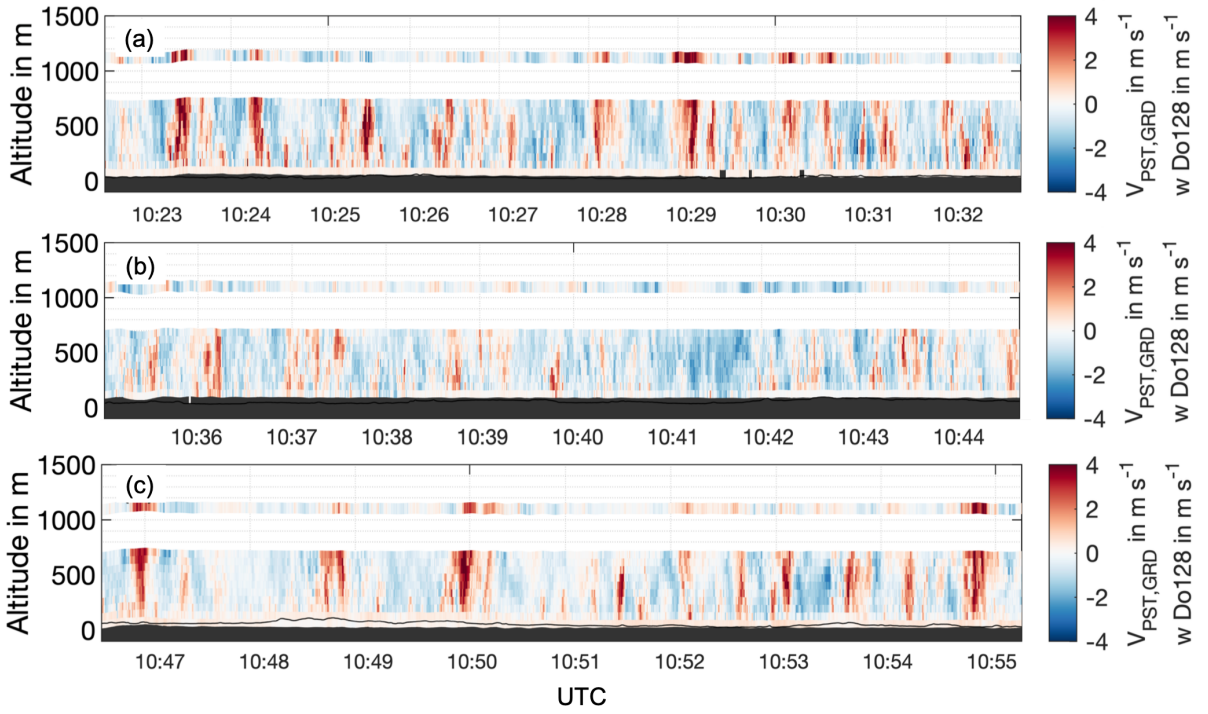

Fig. 6.29.: ADL retrieved and Do128 in-situ vertical wind for three nadir transects during the flight on 19 July. (a) first across-wind transect, (b) along-wind transect (upwind) and (c) second across-wind transect.

6.7. Application 3 - Differences between along- and across-wind turbulence probing

The third application study utilizes two across-wind and one along-wind transect (upwind) obtained during the flight on 19 July 2019 to illustrate differences in turbulence characteristics between along- and across-wind sampling. All transects were conducted at a flight level of 1100 m below a cloud base at approx. 1300 m. During the transects an average wind speed of 5 $\mathrm{m\,s^{-1}}$ at flight level and 8-10 $\mathrm{m\,s^{-1}}$ above the BL prevailed (for a description of the meteorological conditions during the flight see Sec. 6.1).

Fig. 6.29 shows the motion and wind corrected ADL and Do128 in-situ measurements for the three transects. As expected, differences are already discernible with the naked eye. Whereas the across-wind transects exhibit a semi-regular spacing of strong updrafts alternating with wider but weaker intermediate downdrafts, the along-wind transect shows a different picture. The along-wind transect was conducted in an extended region of sinking motion, with one particularly noticeable stretch from 10:40-10:43 UTC where blue colors prevail (corresponding to 12 km distance). Updrafts are less clearly defined, less regular and less intense compared to the across-wind transects. No strong updrafts reach the flight level. In contrast, for the across-wind

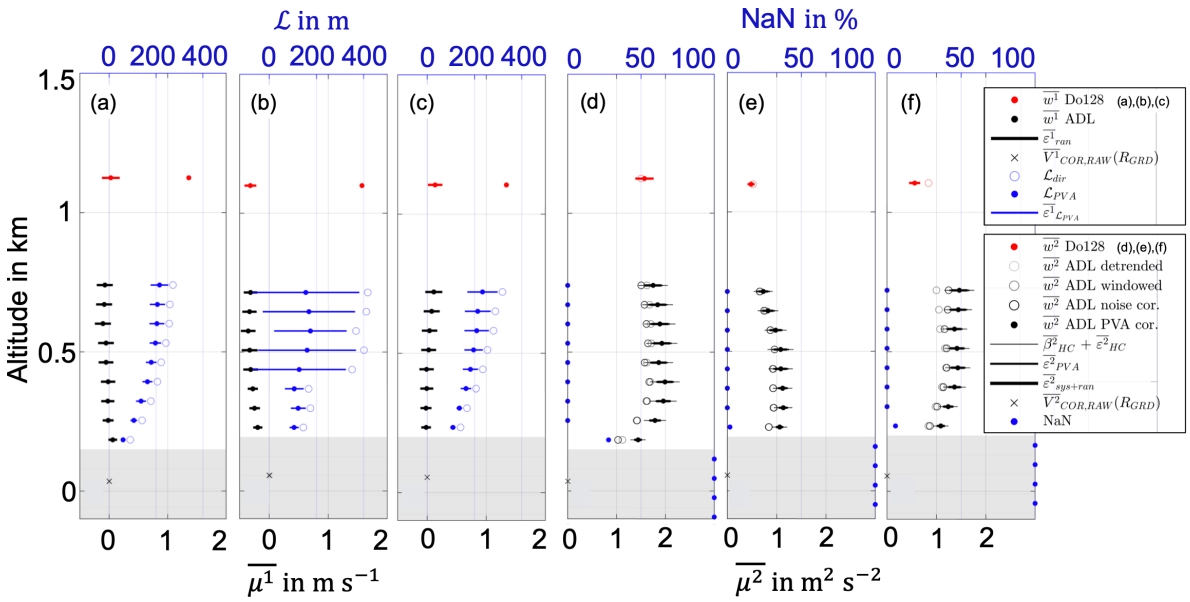

Fig. 6.30.: Profiles of statistical parameters associated with the three transects in Fig. 6.29. (a), (b), (c) mean vertical wind and integral length scales, (d), (e), (f) vertical wind variance and fraction of invalid data. Display characteristics as in Fig. 6.25.

transects some of the strongest updrafts do reach the flight level as measured by the Do128 in-situ. While an intensity reduction of intermediate strength updrafts is detectable towards the top, this is not the case for high intensity updrafts which protrude upward with undiminished strength. These updrafts gain intensity from the middle of the BL towards the top and appear to widen, although the latter is difficult to diagnose based on the two-dimensional cross-sections at hand. This behavior can be explained by the development of clouds above the strongest updrafts, which support updraft strength through latent heat release. Weaker updrafts, which do not reach the BL top, lack the latent heat invigoration.

The turbulence characteristics observed from the time series are also reflected in the vertical profiles (Fig. 6.30). Whereas the two across-wind transects exhibit a mean vertical wind close to zero, the along-wind transect shows consistently negative values throughout the BL, with the magnitude decreasing towards the surface. Associated with the increasing updraft width, the integral length scale increases from the surface towards BL top. Extrapolating the ADL retrieved values yields good agreement with the Do128 in-situ measured ones. The along-wind PVA corrected integral scales are smaller than the across-wind ones, pointing to the anisotropy of the observed turbulence (Lenschow and Stankov, 1986). An interesting phenomenon occurs for the direct integral length scales, calculated for the along-wind transect above the lowest three levels. A sudden, unrealistic jump in the direct integral length scale estimated is observed, whereas the von-Karman model PVA corrected values follow a more smooth profile. This jump

illustrates the problem of calculating a reliable integral length scale in some conditions. The jump is caused by slightly positive autocorrelation over many legs, which is interrupted by a short zero crossing at the beginning for the lowest altitudes, whereas no zero-crossing occurs at higher altitudes (not shown). Due to the integration of the autocorrelation function up to first zero crossing large differences can occur. Here, the fit of an exponential function as conducted by Lothon et al. (2009) might be more meaningful for the non-PVA corrected integral length scale estimate. For now, the discrepancy is discussed for illustrative purpose and to show that the von-Karman PVA correction provides more reasonable estimates in this case (as it is based on a fit to the structure function) and captures the associated ambiguity by showing an increased uncertainty.

The differences in turbulence characteristics are also visible in the vertical wind variance profiles (Fig. 6.30d, e, f). Clearly, across-wind transects show higher turbulence intensity than the along-wind transect. Whereas the values at the lowest level are still somewhat similar with values around $1\,\mathrm{m^2\,s^{-2}}$, the variance profile for the along-wind transect then shows a noticeable decrease towards BL top, which can be explained by the large scale subsidence observed on the along-wind leg. The two across-wind profiles remain near-constant with altitude in the middle BL, which is in-line with other observed profiles in the cloud-topped convective BL (Lenschow et al., 2012; Kalthoff et al., 2013). Noticeably, the last across-wind transect also shows a decreased in-situ measured value of vertical wind variance at approx. $0.9z_i$. Only a few strong updrafts protrude to flight level and calmer conditions prevail in between, which illustrates the importance of entrainment and detrainment processes at the BL top.

Last, spectral analysis also presents a picture consistent with the previous observations (Fig. 6.31). Associated with the more intense and semi-regular updrafts, the across-wind transects show higher spectral intensities in the production range than the along-wind transect. Further, the across-wind transects show a defined peak and rapid decay towards the spectral gap, whereas this is not the case for the along-wind spectra (e.g. due to the extended downdraft section). All spectra show a shape in accordance with the $-2/3$ decay expected in the inertial subrange. As discussed before, a noticeably steeper decay occurs for wavelengths smaller than approx. 350 m due to the PVA effect of the lidar. Similarly, the uptick in ADL spectral levels for highest measurement frequencies, due to uncorrelated noise, is also present again. Especially the across-wind spectra show a distinct shift of the spectral peak with altitude. In agreement with theory (Kristensen et al., 1989) and previous studies (Kaimal et al., 1976), closer to the ground the peak is located at shorter wavelengths whereas it is shifted towards longer wavelengths at the top of the BL. Good agreement is observed between the Do128 in-situ measured spectra and the ADL retrieved ones, especially for the shape and course of the spectra, whereas the levels

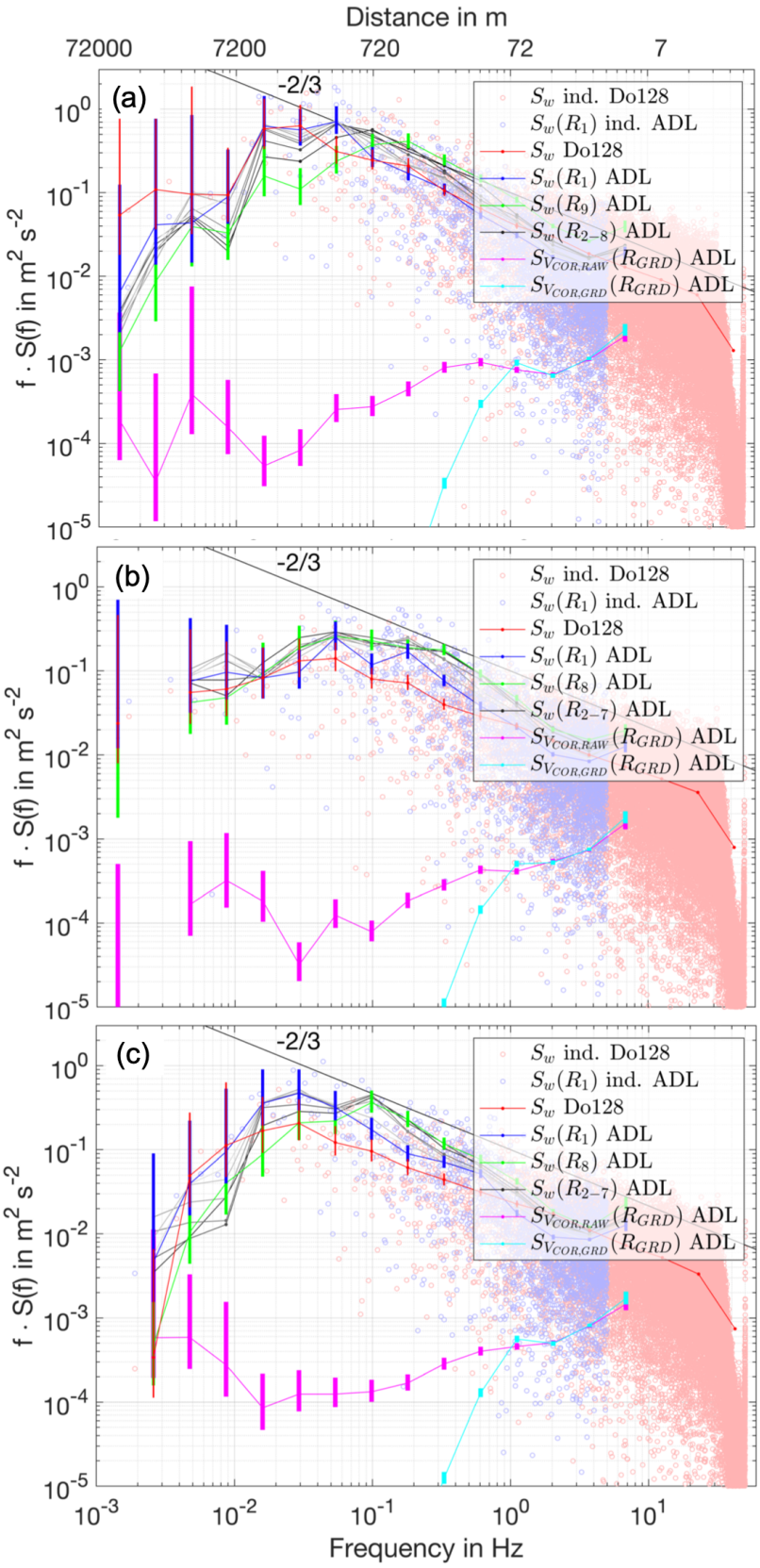

Fig. 6.31.: Spectral decomposition of the vertical wind time series shown in Fig. 6.29. Display characteristics as in Fig. 6.26.

differ slightly at flight level in accordance with the observed variance profiles (Fig. 6.30). The spectra of the measured ground return velocities show low levels of variance and little variation with frequency, proving once again the high accuracy of the motion correction procedure during straight and level flight.

Overall, the results and discussion illustrate two points. First, the prototype ADL measurements allow for previously unavailable insight into BL turbulence with an unprecedented level of detail and flexibility. Second, ground-based measurements, which have to rely on the advection of turbulence with the mean wind, may struggle to capture BL characteristics adequately. Assuming Taylor's hypothesis of frozen turbulence, the observed along-wind transect requires approx. two hours for advection past a ground-based measurement station at the observed mean wind speed. Even if the air mass does not change it's characteristics during this time period, which is unlikely, the turbulent exchange processes in the air mass can hardly be assumed to be a representative sample of the BL state over an extended region, as revealed by the across-wind transects.

7. Conclusion and Outlook

There is a strong need for further improvement of observational capabilities with respect to wind measurements, especially in the turbulent BL. This work is dedicated to the commissioning of a new ADL system, which provides measurements of the turbulent wind field inside the BL at highest resolution up to date. Due to the turbulent measurement environment and high resolution, previously unaddressed challenges mandate investigation, both for AVAD wind profiling and nadir measurement techniques. To overcome these challenges, a two-step approach based on virtual observations and real-world prototype measurements is pursued. Hence, this work is able to provide a previously unavailable full assessment of the potential of ADL measurements for BL research. Thereby, the achievable high measurement accuracy is proven and the added value of ADL measurements for BL research is demonstrated.

7.1. Advancing ADL wind measurements for BL research using virtual observations

As a first part of this work a novel LES-based airborne Doppler lidar simulator is developed. The ADLS utilizes LES generated wind fields at 10-m grid spacing to simulate the turbulent convective BL. To capture the variability of BL turbulence, two different LES sets A and B are utilized which are driven by different surface sensible heat fluxes, approx. $170\,\mathrm{W\,m^{-2}}$ for set A and $30\,\mathrm{W\,m^{-2}}$ for set B. LES set A employs five background wind cases at $0, 2, 5, 10$ and $15\,\mathrm{m\,s^{-1}}$ and LES set B three background wind cases at $5, 10$ and $15\,\mathrm{m\,s^{-1}}$. The emulated measurement system consists of an aircraft, scanner and lidar, considering all geometric transformations applicable in real-world measurements. After system emulation, a simulated measurement is conducted based on the LES wind fields. The ADLS does neither include a simulation of the signal return (aerosol scattering) processes nor of the lidar instrument physics. These processes have been investigated adequately elsewhere and problems associated with them are shown to be of manageable magnitude in practice. Instead, a direct measurement approach using an idealized system is chosen here. The ADLS allows for theoretical insight into the measurement process because the input truth used to create the measurements is known exactly, thereby representation and co-location problems present in real-world studies can be avoided. Hence, a quantitative

analysis of ADL retrieval error due to the violation of the flow homogeneity assumption used in AVAD wind profiling can be provided.

The impact of the violated flow homogeneity assumption on AVAD wind profiling error [1]

The retrieval error is first assessed for a standard system setup and retrieval strategy, based on a most commonly used ADL system setup and retrieval strategy. The system utilizes scans at 60° elevation angle (30° off-nadir) and a scan speed of $18\,^{\circ}\,s^{-1}$, together with a lidar measurement frequency of 10 Hz. AVAD retrieval is conducted using a volume-based technique, with an along-track averaging distance of 1300 m and the same across-track averaging distance. The ADLS simulation enables retrieval of 1680 wind profiles from LES set A and 336 wind profiles from LES set B, more than what is typically available in real-world accuracy evaluations. Results show that ADL wind profiling in inhomogeneous flow is possible with acceptable accuracy, despite the violation of the flow homogeneity assumption used in AVAD. Nevertheless, retrieval errors due to the violation of the flow homogeneity assumption are non-negligible and can present a major contribution to overall wind profiling error. As expected, there is a strong influence of BL turbulence on wind profiling accuracy. Wind profiling in LES set A exhibits an MAE of $0.75\,m\,s^{-1}$ and bias of $0.19\,m\,s^{-1}$ for wind speed, whereas a reduced MAE of $0.29\,m\,s^{-1}$ and bias of $0.05\,m\,s^{-1}$ are present in LES set B.

An important new finding of this study is that Doppler lidar wind profiling can be unreliable and even biased at low wind speeds under turbulent conditions. For low wind speeds, erroneous mapping of the vertical wind into horizontal wind can occur, thereby biasing the retrieved wind speed. The bias is only present in the retrieved wind speed while the horizontal wind components remain unbiased (as a systematically too large wind vector with random orientation is retrieved). Results show that the usage of the condition number CN as a quality control criteria is necessary and adequate, as it can improve wind profiling quality by flagging unreliable measurements (e.g. due to clouds blocking sectors of the scan). However, quality control of the LSQ-fit by the coefficient of determination R^2 can introduce or worsen the discussed retrieval bias for problematic system setups or retrieval strategies.

[1] The results of this section are presented in the publication: Gasch, P., Wieser, A., Lundquist, J. K. and Kalthoff, N., 2020: An LES-based airborne Doppler lidar simulator and its application to wind profiling in inhomogeneous flow conditions. Atmospheric Measurement Techniques, 13(3), 1609-1631.

Measurement system setup and retrieval strategy optimization

In order to minimize the retrieval error due to flow inhomogeneity a measurement system setup and retrieval strategy optimization is conducted using the ADLS. This optimization is done in preparation of an upcoming ADL system for BL research which features flexible scan geometries. While the specific level of wind profiling error depends on the turbulent conditions present, the ADLS allows for insight into system setup and retrieval strategy characteristics which hold generally under turbulent conditions.

An adequate system setup improves wind profiling quality. Scan elevation angles steeper than 60° are problematic under turbulent conditions in the BL, due to the strong influence of vertical wind on radial velocity and thereby retrieval error. For elevations greater than 60°, the wind profiling error grows rapidly and the retrieval becomes increasingly biased at low wind speeds. A faster scan speed decreases retrieval error, if random system noise remains low. Fast scans decrease correlation between adjacent measurements, thereby providing more information for the wind profile retrieval. For the same reason, an increase in measurement frequency is not beneficial for retrieval accuracy because the additional measurements obtained are highly correlated, providing no additional information.

Retrieval settings also impact wind profiling quality. As expected, longer horizontal averaging distances increase retrieval accuracy. However, the retrieval quality from short averaging distances can be quality controlled using the reliable CN criterion. Therefore the increase in wind profiling error is below linear for short averaging distances. In comparison, the vertical averaging distance does not exhibit a strong impact on retrieval error. Vertical averaging improves retrieval quality only marginally, due to the similar structure of turbulence over multiple range gates. Under the conditions investigated here, the wind profile retrieval is robust if at least 180° of the azimuth scan sector are available, at slightly increased error levels which can again be reliably controlled using the CN.

Overall, the ADLS results allow for determination of preferential wind profiling strategies to achieve a confined measurement footprint, while maintaining high accuracy and resolution. As a demonstration, an optimized system for BL wind profiling is presented. The system utilizes scans with a shallow 45° elevation angle and a fast scan speed of $45\,^{\circ}\,\mathrm{s}^{-1}$, together with a high lidar measurement frequency of 10 Hz. The resulting increase in across-track measurement footprint to 2200 m is offset by high-resolution retrieval volumes confined to 325 m in along-track direction (using forward and aft looking lidar shots from aircraft positions before and behind the retrieval volume). With such a setup, flow phenomena with one symmetry axis can be investigated at highest resolution when the aircraft is flying perpendicular to the flow

phenomenon under investigation. The increase in retrieval quality is powerful. Despite the 30% reduced measurement footprint, the wind profiling error decreases by approx. 30% due to the shallow scans. Further, five times more wind profile points can be retrieved due to the increased along-track and vertical resolution compared to the standard system setup and retrieval strategy.

Data-driven uncertainty assessment of the retrieved wind profiles

To generalize the ADLS results and enable transfer to real-world measurements the applicability and reliability of a data-driven uncertainty estimation method is vetted as a final step of the theoretical study. In order to achieve an uncertainty estimation, flow inhomogeneity leading to profiling error is separated into sub-scan-volume and supra-scan-volume contributions.
Results show that the residual variance, obtained from the AVAD LSQ-fit, is a useful estimator of the unknown data variance, if corrected for correlation among the residuals. Propagation of the residual variance through the covariance matrix leads to a conservative estimation of the wind profiling uncertainty due to sub-scan-volume flow inhomogeneity. Uncertainty due to supra-scan-volume flow inhomogeneity cannot be captured by the residual variance, as it results in erroneous mapping and is thereby not detectable in the residuals. Therefore, a parameterization of the supra-scan-volume contribution is developed, which draws on basic available BL parameters (system and retrieval parameters, BL height, average updraft vertical winds).
The reliability of the data-driven uncertainty estimation method is systematically validated for a range of atmospheric conditions, varying system setups and retrieval strategies. Analysis of the mean estimated uncertainty and statistical properties of the estimation method proves that the proposed method works reliably and yields a conservative uncertainty estimation. It is therefore suggested that this method be included in standard retrieval schemes used for AVAD retrieval, besides the usually conducted validation by other measurements systems.

7.2. Advancing ADL wind measurements for BL research using a prototype system

The second part of this works analyzes real-world ADL measurements from four research flights conducted in the vicinity of Brunswick in July 2019 and applies the ADLS developed concepts. Although only a fixed-beam prototype system without scanning capabilities was available for the measurements, the conducted measurements allow for a full assessment of the potential of ADL measurements for BL research (e.g. wind profiling is conducted by banking the aircraft in turns). Thereby, the up to date highest resolution nadir measurements of the vertical wind

inside the turbulent BL can be presented, which provide previously unavailable insight into BL flow phenomena and turbulence characteristics.

Data processing and quality control scheme

In order to achieve data synchronization, calibration and correction as well as an accuracy assessment the procedures of multiple existing studies are combined, refined and extended as a part of this work. Hence, a new level of measurement quality evaluation for ADL systems in the turbulent BL is attained.

Time synchronization is carried out using a new procedure, which exploits the variation in the maximum lidar sensing distance (e.g. distance to the ground) during aircraft roll maneuvers. Thus, a time synchronization with higher accuracy than the 0.1 s lidar measurement resolution is obtained. A beam pointing-angle calibration is conducted following the techniques outlined in Haimov and Rodi (2013). The beam pointing-angle calibration yields reliable results, although problems with the accuracy of the aircraft INS during turn maneuvers are revealed. The lidar radial velocity noise is analyzed using methods proposed by Lenschow et al. (2000). Low levels of radial velocity noise are diagnosed for the WTX system, although a high measurement frequency of 10 Hz is utilized. The error of the INS based motion correction is assessed through a statistical analysis of the measured ground return velocities. A reliable ground return identification allows for detailed insight into the error associated with the motion correction procedure, including a classification by flight state (distinguishing between nadir, turn and PRY maneuvers). Thus, reliability and accuracy of the motion correction procedure using the aircraft INS data is proven. In a further step, a refined motion correction procedure is developed which corrects for remaining error in the INS motion correction based on the residual ground return velocities. In doing so, a further strong improvement in motion correction accuracy can be demonstrated.

For the retrieval of statistical properties of the vertical wind in the turbulent BL, such as the mean and variance, further effects require consideration. The contribution of the horizontal wind due to non-nadir beam pointing is removed following a procedure applied by Chouza et al. (2016b). In this, the horizontal wind contribution is removed by using the horizontal wind profiles obtained from ADL wind profiling conducted before and after each nadir transect. The remaining uncertainties due to non-nadir beam pointing and horizontal wind contamination are quantified based on the theory outlined by Strauss et al. (2015). Thereby it is shown that a reliable retrieval of the mean vertical wind is possible, especially when using the refined motion correction procedure. Retrieval of the vertical wind variance requires further consideration

of the PVA effect. A modification of the von-Karman turbulence model founded technique, developed by Frehlich (1997), is implemented to correct the effect of PVA on the measured variance and integral length scales. Applying evaluation schemes suggested by Brugger (2014) a reliable functionality of the PVA correction procedure is demonstrated. At last, the statistical uncertainties due to the limited measurement transect lengths are also taken into account following Lenschow and Stankov (1986).
Overall, the data processing and quality control results show that reliable measurements are possible even when the aircraft is flying inside the turbulent BL, which is a challenging measurement environment due to the turbulence disturbing aircraft flight and motion. It is recommended that future ADL studies apply a similar scheme for measurement accuracy assessment in order to represent all relevant errors and quantify the associated uncertainties.

Accuracy estimates for AVAD wind profiling and nadir measurements

The in-depth data processing scheme, together with the ADLS developed methodology, allows for an evaluation of both the AVAD wind profiling and nadir measurement quality.
Raw measurement system errors are of manageable magnitude and can be controlled using the measured ground return velocities for measurements of the mean vertical wind and wind profiling purposes (Tab. 6.3). Thereby, for wind profile retrieval the violation of the flow homogeneity assumption is left as the major driver of wind profiling error, confirming the assumptions presented in the ADLS. Hence, the ADLS as an idealized measurement system test bed and the data-driven uncertainty estimation method developed therein prove their usefulness, as does the ground return based radial velocity accuracy assessment of the real ADL system. For real-world comparisons of ADL retrieved wind profiles, additional problems due to co-location, timing and errors in the reference measurement system cause considerable differences, which are not attributable to the ADL wind profiling accuracy. Nevertheless, a comparison of the ADL retrieved and Do128 in-situ measured wind profiles (obtained from wind profiling in close proximity) for 900 wind profile points yields good agreement. Conservative error bounds are obtained with an MAE of $0.58\,\mathrm{m\,s^{-1}}$ and a bias of $0.01\,\mathrm{m\,s^{-1}}$ for wind speed, as well as an MAE of $7°$ and a bias of $1°$ for wind direction. An increase in error with increasing horizontal distance between the two measurement systems is present, as is an increase of retrieval error with increasing turbulence intensity. The effective accuracy of the ADL retrieved wind profiles can be higher than given by the MAE when using the measurement situation dependent data-driven uncertainty estimation method (e.g. for low turbulence situations).

The reliable retrieval of wind profiles enables removal of the horizontal wind contribution and thus retrieval of the vertical wind with high accuracy. Another highly conservative error estimate is obtained through a comparison of the mean vertical wind obtained on extended legs to a $0\,\mathrm{m\,s^{-1}}$ reference value mandated by continuity. Thereby, the benefit of the refined motion correction scheme can be demonstrated, which yields a bias $< 0.03\,\mathrm{m\,s^{-1}}$ and a random uncertainty of approximately $0.13\,\mathrm{m\,s^{-1}}$ for vertical wind measurements.
Enabled by the slow flying aircraft and high lidar measurement frequency, this work presents nadir measurements of the vertical wind inside the BL at previously unavailable resolution. Thus, ADL measurements of the vertical wind variance inside the BL are obtained, which rely on the thorough evaluation of systematic offsets and random uncertainties, as reference measurements are unavailable. To enable accuracy assessment, theoretical models and methods from a number of studies named above are combined, refined and extended. All applied models and methods allow for estimation of the associated offsets and uncertainties on an individual leg and range gate basis. Hence, this works presents the most comprehensive accuracy assessment of ADL variance measurements up to date (Tab. 6.4). Overall, for reliable variance measurements, corrections of the systematic effects of random radial velocity noise, PVA and underestimation due to finite measurement lengths are shown to be important, as is the estimation of the systematic variance contamination due to motion correction and by the horizontal wind, associated with non-nadir beam pointing. Random uncertainties due to horizontal wind contamination and finite measurement lengths also require consideration and can be of non-negligible magnitude. Exemplary agreement between Do128 and ADL variance measurements is shown outside the BL. Another case, in which the lidar path was retraced by the Do128 in close spatial proximity after completion of an initial leg, also shows good agreement.
Overall, building on the comprehensive uncertainty investigation as well as the few available in-situ comparisons, the ADL measured variance is regarded to be accurate and reliable after application of all post-processing procedures, within the bounds of the associated uncertainty estimation. Thus, it is assessed that insight into BL turbulence characteristics can be gained by the conducted ADL measurements, offering exciting new possibilities for BL research.

The added value of high resolution ADL BL measurements

Based on these results, three exemplary application studies are analyzed from the conducted test flights. By this, the new insights and added value which can be provided by a high-resolution ADL system are demonstrated.

In the first application study, the prototype ADL is used for a demonstrative retrieval of vertical wind statistics in a clear section of a turbulent BL topped with shallow cumulus humilis elsewhere. Added value is provided as the ADL is able to retrieve the vertical wind statistics at all levels simultaneously with reasonable uncertainty. The thereby observed variance profile is in good agreement with previously reported profiles (Lenschow et al., 1980), with a maximum in the lower half of the BL and a decrease towards the top. Due to the high resolution a spectral analysis of the measurements is possible up to high frequencies. Good representation of the expected spectrum in the production range is observed, with a shift of the spectral maximum from shorter wavelengths at the ground towards longer wavelengths at the top of the BL. Further, agreement with the theoretical -2/3 slope in the inertial subrange is observed for wavelengths > 350 m, whereas for smaller wavelengths the ADL spectra drop off more rapidly due to the PVA effect, as expected.

In a second application study, the potential of the prototype ADL system to provide spatially resolved measurements of both the horizontal and vertical wind in moderately complex terrain is demonstrated. To this end, multiple transects across the Harz mountain range along the same ground path are analyzed from a morning and an afternoon flight. Thereby, both the temporal and spatial development of the BL, as well as its link to orographic features, can be captured. Transects acquired in the morning show a distinct difference in BL structure between the orographically structured mountain range and its surroundings. Above the mountains, the BL reaches higher and development has progressed further, stronger turbulence compared to the surroundings is noticeable. Rapid temporal development of the BL is observed between transects. In the afternoon, the scene is dominated by the presence of convective clouds and a region of higher wind speeds passing the mountain range. The BL height is less influenced by the mountain range, with only a slight increase which does not reflect the increased terrain height in full. The observed vertical wind scene is heterogeneous. Compared to the morning, much stronger, wider and deeper updrafts dominate the scene, which are connected to the cumulus clouds present. Above the main Harz mountain ridge, a more 'fuzzy' updraft scene can be detected, with horizontally more narrow but nevertheless strong updrafts, spaced closer together. In comparison, areas in the direct vicinity of the Harz show extended areas of subsidence and more narrow and shallow updrafts. Potentially, the change in BL structure is connected to the stronger horizontal across-ridge flow and/or a superimposed mountain venting circulation (De Wekker and Kossmann, 2015). Further studies and combination with ground-based measurements are needed to trace and distinguish these processes above the mountainous terrain better.

The Harz transect example demonstrates the added value of an ADL compared to ground-based systems. Through the moving platform, the spatial variability of the wind field is captured, providing information which is otherwise inaccessible, but which is key to understanding the full picture. Synergies can be achieved by combining an ADL system with ground-based observations. While the former is able to capture the spatial variability, the latter are able to capture the temporal evolution at a fixed point. The combination of both measurement systems can provide valuable new insights into BL properties, especially over complex terrain.
A third application study analyzes differences in BL turbulence characteristics observed between across- and along-wind sampling. All transects show consistent characteristics between the ADL retrieved vertical winds and the Do128 in-situ measurements 400 m above for individual updrafts, leg averaged mean values and spectra. Across-wind turbulence properties are dominated by the semi-regular spacing of strong and narrow updrafts, with weaker but wider intermediate downdrafts. Analysis of both profiles and spectra reveals characteristic BL turbulence phenomena, e.g. an increasing organization of updrafts towards the BL top and convective invigoration of updraft strength through latent heat release in cumulus clouds. The along-wind transect on the other hand is conducted in an extended meso-scale region of descending air. It thereby exhibits much weaker turbulence with shorter integral length scales. The altered turbulence characteristics in along-wind sampling illustrate the potential problems of ground-based measurement systems (and thereby the advantage of an ADL). Ground-based measurements have to rely on the mean wind for advection, which can be problematic due to the observed organization of turbulence in along-wind direction.

7.3. Outlook

Overall, this work highlights the potential of ADL measurements for BL research. Both wind profiling and nadir measurements are possible with high resolution and accuracy, thus allowing for previously unavailable insight into BL flow characteristics with an unprecedented level of detail. Further, the benefits of LES-based ADL simulations are demonstrated. Linking theory and measurement, they provide a valuable and cost-effective tool to optimize the quality of ADL measurement results. The ADLS offers a promising opportunity to investigate the ADL measurement process more closely, not only with respect to wind profiling, but also with respect to retrieving turbulent quantities from nadir measurements. For example, the methods proposed by Strauss et al. (2015) have not been vetted systematically for validity and possible improvement, which can be done using an ADLS. Further, the ADLS can also serve as an algorithm testbed for more complex retrieval techniques such as airborne dual-Doppler.

As mentioned in the beginning, the full scanning measurement system is available with the completion of this work. It provides for exciting new possibilities and can present another step towards new insight into BL turbulence. Extending the capabilities of the prototype system, the full system is able to provide both nadir and wind profiling measurements at highest resolution (using the ADLS developed optimized system setup and retrieval strategy). Associated with the new system two measurement campaigns are planned already. The first focuses on the downstream wake effects of wind farms above the North Sea (Platis et al., 2018) and uses the ADL to provide two-dimensional insight into flow phenomena. The second focuses on the characteristics of convective system across the Swabian Alps. Here, the ADL is used to gain insight into the meso-scale flow phenomena and turbulence characteristics in the vicinity of these convective systems. Both measurement campaigns will greatly benefit from the high resolution measurement strategies developed within this work. Finally, ADL capabilities are of great importance in future research activities in mountainous terrain, where complex flow is ubiquitous and can only be assessed in limited areas using ground-based measurements alone (Adler et al., 2020b). While early studies already used coarse ADL capabilities (Weissmann et al., 2005a), high resolution systems can enable a new level of insight and provide valuable information in large scale experiments such as TeamX (Rotach et al., 2020).
Due to the complexity of BL characteristics and flow, a further development of airborne observational capabilities is advised. In this respect, the further advancement of ADL capabilities and implementation of sophisticated scanning strategies such as airborne dual-Doppler appears promising. Further miniaturization of laser and scanner components could allow implementation of ADL systems on smaller aircraft platforms (e.g. motor gliders), which would enable more widespread and low-cost measurements. Besides the development of pure ADL capabilities, a further extension and combination of ADL measurements with other measurement systems is also worthwhile. In addition to the invaluable in-situ turbulence probing equipment, other desirable measurement capabilities include radiation sensors as well as nadir and zenith directed cameras to capture the location of clouds with respect to the position of the lidar beam. The combination of ADL with other lidar systems, e.g. DIAL for humidity profiling, is neither low-cost nor low-complexity but definitely worth pursuing nevertheless, as it allows for retrieval of the important BL turbulent fluxes (Kiemle et al., 2007, 2011). Of course, combination of ADL with ADR measurements is also promising, as it enables continuous retrieval of the wind field outside and inside of clouds. However, due to weight and balance as well as power limitations it appears difficult to achieve based on one aircraft, therefore simultaneous coordinated aircraft measurements offer a way forward in this direction. Similarly, coordinated parallel in-situ measurements by a second aircraft also offer a way forward to validate

the uncertainty estimation concepts and theories applied in this work (although difficult and expensive to achieve). Especially a validation of the horizontal wind contamination uncertainty estimation and the PVA correction procedure seems worthwhile to conduct for ADL variance measurements and has not been attempted to my knowledge. Last, ADL measurements in field campaigns should be conducted in close coordination with ground-based measurements. The synergy of spatially resolved, high resolution ADL measurements in combination with extensive ground-based measurements, covering the temporal evolution over extended periods, is certain to provide new and valuable insight into atmospheric phenomena.

A. Additional concepts and theory

A.1. Definition of the terms error, uncertainty and accuracy

This work employs a distinction between error, uncertainty and accuracy based on ISO 5725-1:1994 and JCGM 100:2008.
The term error refers to (observed or expected) quantifiable deviations compared to a reference, which is either the truth or treated as the truth. The overall error includes contributions due to systematic error (also called bias) and random error following ISO 5725-1:1994 . The actual error can only be determined if the input truth is available for comparison (e.g. only in the ADLS where the input LES wind fields are known, never in real-world situations). Nevertheless, the error can also be conservatively bounded or estimated through a comparison to other reference measurements treated as the truth (e.g. in real-world situations). As these reference measurements and the comparison are subject to errors themselves, a conservative estimate or upper bound of the error can be obtained. In this case, however, it has to be ensured that the error is assessed under representative conditions.
If an error is expected to be present but cannot be quantified by comparison to a reference truth, the estimation of an uncertainty is required for measurement accuracy assessment. The term uncertainty refers to an estimation of the average expected error based on theoretical concepts, here this work follows the guidelines outlined in JCGM 100:2008. Thereby, uncertainty is always associated with estimation methods, which attempt to quantify the average expected error of the measurement process, but without having a reference truth available for error assessment. The uncertainty estimation methods can be data-driven or based purely on theoretical considerations and models. While the uncertainty is thereby similar to an error, as it tries to represent the expected average error magnitude (if the input truth was known or a reference measurement was available), the distinction is nevertheless made to stress the different origins of the estimated quantities. Uncertainty contributions again consist of systematic offsets and random uncertainties. Systematic offsets present an estimation of the expected systematic error magnitude of the measurement. Systematic offsets can in some cases be further corrected and thereby reduced based on theoretical concepts. An example is the underestimation of the measured variance due to pulse volume averaging by the lidar. The random uncertainties present an

estimation of the expected random error magnitude based on theoretical concepts. An example is the random sampling uncertainty due to finite measurement lengths. Random uncertainties are also associated the correction of the systematic offsets.
Accuracy as used in this work refers to any observed or estimated deviation, whether it is obtained from an error or uncertainty assessment, it thereby includes both systematic and random contributions for both measures.
At the given definition, the AVAD wind profiling retrieval error due to flow homogeneity violation is directly accessible in the ADLS as the LES input truth is known. In real-world measurements, a wind profiling error can also be conservatively estimated in comparison to other reference measurements (e.g. dropsonde or aircraft), however, in this case the observed error consists of spatio-temporal co-location errors, measurements system errors (both ADL and reference) and ADL retrieval errors, which cannot be separated.
Real-world ADL measurements of other quantities such as the vertical wind mean and variance rarely offer direct access to an input truth or comparison with reference measurements, therefore they mostly rely on the estimation of uncertainties based on the measured data alone or theoretical concepts. Three exceptions are present in this work: First, the combined error due the time synchronization, beam pointing-angle and motion correction procedures can be conservatively assessed based on the measured ground return velocities. Second, the error due to the horizontal wind removal procedure (as well as the previous three procedures), can be conservatively estimated for the mean vertical wind by comparison to a $0\,\mathrm{m\,s^{-1}}$ reference truth mandated by theory. Third, a conservative assessment of wind profiling error is obtained in comparison to the Do128 in-situ measurements discussed above.

A.2. Coordinate transforms

As outlined in Lenschow (1972) and Leon and Vali (1998), transformations between the LES earthbound (superscript E) coordinate system, oriented east-north-up, and the aircraft (superscript A) coordinate system, oriented along aircraft-right wing-down, 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

$\mathbb{T}^{AE}$ (ARD to ENU). Here, the transformation is 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

$$\boldsymbol{b}^E = \mathbb{T}^{AE}\boldsymbol{b}^A \tag{A.1}$$

and

$$\boldsymbol{b}^A = \mathbb{T}'^{AE}\boldsymbol{b}^E. \tag{A.2}$$

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

$$\mathbb{T}^{AE} = \mathbb{HPR}. \tag{A.3}$$

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:

$$\mathbb{H} = \begin{pmatrix} \sin(\psi) & \cos(\psi) & 0 \\ \cos(\psi) & -\sin(\psi) & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{A.4}$$

$$\mathbb{P} = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix}, \tag{A.5}$$

$$\mathbb{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi). \end{pmatrix}. \tag{A.6}$$

Combined, this results in

$$\mathbb{T}^{AE} = \begin{pmatrix} \sin(\psi)\cos(\theta) & \cos(\psi)\cos(\phi) + \sin(\psi)\sin(\theta)\sin(\phi) & -\cos(\psi)\sin(\phi) + \sin(\psi)\sin(\theta)\cos(\phi) \\ \cos(\psi)\cos(\theta) & -\sin(\psi)\cos(\phi) + \cos(\psi)\sin(\theta)\sin(\phi) & \sin(\psi)\sin(\phi) + \cos(\psi)\sin(\theta)\cos(\phi) \\ \sin(\theta) & -\cos(\theta)\sin(\phi) & -\cos(\theta)\cos(\phi) \end{pmatrix}. \tag{A.7}$$

Additional coordinate transforms for real-world measurements

For real-world measurements, the data processing and retrieval require an additional coordinate system which is fixed with respect to the Earth's center. The aircraft INS provides its position data in reference to the geodetic latitude φ_g, longitude λ_g and height h_g in a geodetic coordinate system (superscript G), using the World Geodetic System WGS84 ellipsoid (NIMA, 2000) for reference. The transformation from GEO to a local ENU coordinate system involves an intermediate Earth-centered, Earth-fixed (superscript C) coordinate system to obtain cartesian coordinates (X, Y, Z).

At first, the geodetic coordinates $\varphi_g, \lambda_g, h_g$ are converted to cartesian ECEF coordinates using the transformation (Hofmann-Wellenhof et al., 2001):

$$X = \left(J(\varphi_g) + h_g\right)\cos(\varphi_g)\cos(\lambda_g) \tag{A.8}$$

$$Y = \left(J(\varphi_g) + h_g\right)\cos(\varphi_g)\sin(\lambda_g) \tag{A.9}$$

$$Z = \left(\frac{b^2}{a^2}J(\varphi_g) + h_g\right)\sin(\varphi_g), \tag{A.10}$$

with

$$J(\varphi_g) = \frac{a^2}{\sqrt{a^2\cos^2(\varphi_g) + b^2\sin^2(\varphi_g)}}. \tag{A.11}$$

In the above a is the equatorial radius and b the polar radius of the ellipsoid, which are $a = 6378.137\,\text{km}$ and $b = 6356.752\,\text{km}$ for WGS84.

Then, the obtained ECEF coordinates can be transferred to a local ENU coordinate system using the transformation matrix (according to the procedure provided above):

$$\mathbb{T}^{CE} = \begin{pmatrix} -\sin(\lambda_g) & \cos(\lambda_g) & 0 \\ -\cos(\lambda_g)\sin(\varphi_g) & -\sin(\lambda_g)\sin(\varphi_g) & \cos(\varphi_g) \\ \cos(\lambda_g)\cos(\varphi_g) & \sin(\lambda_g)\cos(\varphi_g) & \sin(\varphi_g) \end{pmatrix}. \tag{A.12}$$

The transformation $\mathbb{T}^{CE}$ can be inverted by transposing the transformation matrix, as shown for $\mathbb{T}^{AE}$ above. The return operation for transforming from cartesian to geodetic is more cumbersome. It is therefore not shown for the sake of brevity but can be found in the literature on the topic (Featherstone and Claessens, 2007, and references therein). In this work, both transfor-

mations are combined by using the Matlab functions $geodetic2enu$ and $enu2geodetic$ for the return operation, again while specifying the WGS84 reference ellipsoid.

A.3. Quantification and parameterization of the unresolved supra-scan-volume flow inhomogeneity contribution to wind profiling error

In the following, the contribution of supra-scan-volume flow inhomogeneity to the overall wind profiling error is quantified. This quantification is possible using the ADLS, as it gives access to the underlying atmospheric input to the measurement. From the ADLS, the overall deviations of the atmosphere from the mean state can be calculated for each velocity component (exemplarily conducted for w) as:

$$w^{LES} = \overline{w}^{LES} + w'^{LES}. \tag{A.13}$$

These deviations are the flow inhomogeneity at the location of the measurements. Thereby, they cause the wind profile error due to AVAD assumption violation.
Further, the difference between the retrieved mean wind and the true LES mean wind can be mapped into a scan-volume radial velocity difference. For this, the averaged radial velocities v_D^R from the retrieved wind profile are used. In addition, a second radial velocity is modeled using the LES data, which results from projecting the mean LES wind vector (the true average wind velocity) into radial velocities. This procedure gives the radial velocity measurement that would be recorded if the atmospheric flow was truly homogeneous,

$$v_{COR}^{LES} = \mathbb{G}\overline{\boldsymbol{v}}^{LES}. \tag{A.14}$$

Then, the scan-volume radial velocity difference can be calculated as

$$\Delta v_{COR} = v_{COR}^{LES} - v_{COR}^{R}. \tag{A.15}$$

This radial velocity difference expresses the difference between retrieved wind profile and the LES-truth, projected for the full scan geometry. It can thereby be used to quantify the amount of error that is caused by the large scale flow inhomogeneity on the order of the scan volume and beyond. The inversion of this radial velocity difference gives the component-wise error of the retrieved wind velocity compared to the LES reference:

$$\Delta\boldsymbol{v} = \mathbb{G}^{-g}\Delta v_{COR}. \tag{A.16}$$

The amount of variation which is caused by supra-scan-volume flow inhomogeneity and not detectable through the residual variance is quantifiable through the degree of linear correlation of the component deviations w'_{LES} with the radial velocity difference Δv_{COR}. A supra-scan-volume deviation appears via a linear contribution to the radial velocity difference, as the radial velocity difference is a smooth function on the scale of the scan volume (the scale that matters with respect to supra-scan-volume contributions). Sub-scan-volume contributions to the retrieval error do not contribute to the linear correlation coefficient, due to their shorter scale and thereby noisy character compared to the velocity difference vector, these are expected to be captured through the covariance uncertainty propagation. Therefore, the linear correlation coefficient between the deviations and the scan-volume radial velocity difference is calculated as:

$$\varrho_w = \varrho(w'^{LES}, \Delta v_{COR}) = \frac{\frac{1}{N-1}\sum_{n=1}^{N}(w'^{LES}_n - \overline{w'^{LES}})(\Delta v_{COR_n} - \overline{\Delta v_{COR}})}{\sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(w'^{LES}_n - \overline{w'^{LES}})^2} \cdot \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(\Delta v_{COR_n} - \overline{\Delta v_{COR}})^2}}. \tag{A.17}$$

For complex scan pattern with varying elevation, this calculation must be done component-wise as horizontal wind components can also cause bias (see Koscielny (1984), and Boccippio (1995)). This study deals only with constant elevation scans, therefore the analysis is restricted to the influence of the vertical wind here.

The amount of retrieval error caused by the supra-scan-volume flow inhomogeneity (the non-detectable co-variation of each component with the measurement geometry) is then the scan-volume radial velocity difference projection (the full error, see above) of each component-wise deviation, multiplied with the respective linear correlation coefficient (which expresses the error contributed due to supra-scan-volume flow inhomogeneity of each component),

$$\Delta v^{w'}_{COR} = \mathbb{G}\begin{bmatrix} 0 \\ 0 \\ w'^{LES} \cdot \varrho_w \end{bmatrix}. \tag{A.18}$$

Using this fraction of the radial Doppler velocity, its impact on the retrieved wind vector is obtained using the standard inversion process (Sec. 3.3):

$$\Delta \boldsymbol{v}^{w'} = \mathbb{G}^{-g} \Delta v^{w'}_{COR}. \tag{A.19}$$

To calculate the contribution of this velocity difference vector to the overall retrieval error it remains to calculate its projection onto the overall velocity difference vector:

$$\Delta \boldsymbol{v}_{proj}^{w'} = \mathrm{k}^{w'} \Delta \boldsymbol{v}, \tag{A.20}$$

with

$$\mathrm{k}^{w'} = \frac{\Delta \boldsymbol{v}^{w'} \cdot \Delta \boldsymbol{v}}{\Delta \boldsymbol{v} \cdot \Delta \boldsymbol{v}} = \frac{\Delta \boldsymbol{v}^{w'} \cdot \Delta \boldsymbol{v}}{|\Delta \boldsymbol{v}|^2}. \tag{A.21}$$

Similarly, the retrieval error of the wind speed can be calculated according to

$$\Delta \mathrm{V}_{proj}^{w'} = \mathrm{k}^{w'} \Delta \mathrm{V}_m. \tag{A.22}$$

Thereby, the amount of retrieval error that is caused by supra-scan-volume deviations in the wind field can be expressed. It should be noted that this quantification of error, as opposed to uncertainty estimation, is only possible if the input truth is known (i.e. in the ADLS) and thereby not quantifiable from real-world measurements.

Parameterization of the supra-scan-volume flow inhomogeneity contribution in a data-driven uncertainty estimation method

In real measurements, the error due to supra-scan-volume flow inhomogeneity is unknown because the atmospheric wind vector at the lidar measurement locations is unknown. Further, the wind profiling uncertainty due to supra-scan-volume flow inhomogeneity cannot be quantified using the covariance matrix based approach, because it does not show up in the LSQ-fit residuals. Therefore, the expected error due to supra-scan-volume flow inhomogeneity needs to be parameterized to allow for a data-driven uncertainty estimation.

Two ways of parameterization are possible. If an ADLS is available the average magnitude of mapping can be determined based on its results (e.g. by averaging the $\Delta \mathrm{V}_{proj}^{w'}$ values). This method is very accurate, as it can be tailored to the specific measurement system, retrieval strategy and atmospheric condition. However, two reasons can prevent application of an ADLS. First, it can be very expensive to apply for a variety of flow conditions encountered during one flight and second, many parameters describing the state of the atmosphere in an LES are usually unknown during the flight.

Therefore, a physical parameterization of the expected error is proposed here which is based on a few more easily available boundary layer parameters. The ADLS results allow for a critical

evaluation and validation of the physical parameterization for a wide range of measurement system setup, retrieval strategies and atmospheric conditions. While the physical parameterization is less accurate it nevertheless shows satisfying results and wide applicability.

The critical parameters which define the likelihood and severity of erroneous mapping are the vertical wind magnitude, the scan elevation angle and the ratio of the scan area to the dominant turbulent eddy area, which describes the likelihood of erroneous mapping by supra-scan-volume fluctuations.

The average absolute vertical wind magnitude $\overline{|w(z)|}$ is important as the erroneous mapping is caused by and scales with this quantity. For parameterization, the mean of the absolute vertical wind (vertically resolved) is used based on the LES data. The results from Sec. 6 show that this quantity can be obtained with sufficient accuracy from real-world ADL nadir measurements, which should be used to complement wind profiling.

The average scan elevation angle $\overline{\Phi}$ is an important factor in the parameterization as it determines the projection of vertical wind into horizontal wind. It is directly available from the measurement system.

The (altitude dependent) ratio of the dominant turbulent eddy area A_{eddy} to the scanned area A_{scan} is important, as it expresses the ratio of wind profile error due to sub- compared to supra-volume flow inhomogeneity in a turbulent BL. For a large scanned area, many times the dominant turbulent eddy size, the supra-scan-volume contribution is less crucial as most variations occur on the sub-scan-volume scale and are thereby captured by the residual radial velocity variance. With decreasing scanned area or increasing dominant turbulent eddy size, the importance of supra-scan-volume contributions to wind profiling error increases.

The scanned area is calculated as the product of the altitude dependent scan diameter X_d (determined by the scan elevation angle and the aircraft altitude) and the along-track averaging distance X_a used for retrieval (assuming a rectangular geometry for simplicity). The scan diameter can be calculated from the measurement system setup for any given altitude.

The dominant area contributing to erroneous mapping is calculated by assuming a field of circular eddies consisting of up- and downdrafts. For the dominant turbulent eddy wavelength λ_{max} a value of $\lambda_{max} = 1.5 z_i$ is assumed based on the parameterization by Caughey and Palmer (1979) for the spectral peak wavelength, in-line with the approach by Eberhard et al. (1989). No altitude variation is assumed, as this eddy wavelength can exist throughout the boundary layer and thereby the potential for erroneous mapping is given (although the peak wavelength containing the maximum turbulent energy varies with altitude (Caughey and Palmer, 1979), the decrease in energy for the dominant wavelength is captured through the reduction in vertical wind magnitude). The dominant turbulent eddy wavelength has to be multiplied by 0.5 as the

critical length for erroneous mapping is half the wavelength, e.g. when updraft and downdraft maximum coincide with the scan circle diameter (Fig. 3.4).
At last, due to the circular eddy assumption the areal coverage of critical structures has to be estimated. Here, the results obtained by Lenschow and Stephens (1980) are used which give an areal coverage of 28% for updraft structures, again similar with the approach by Eberhard et al. (1989). This parameter can be refined by using nadir vertical wind measurements of the ADL itself.
The full parameterized uncertainty is multiplied with a factor of 2, because any parameterized offset can be positive or negative, as, unlike for the ADLS-based direct error analysis, the sign of the change is unknown. Overall, we arrive at the following parameterization:

$$\Delta V_{phy}(z) = 2\overline{|w(z)|}\tan(\overline{\Phi})\frac{A_{eddy}}{A_{scan}(z)}0.28 = 2\overline{|w(z)|}\tan(\overline{\Phi})\frac{0.25\pi\lambda_{max}^2}{X_d(z)\cdot X_a}0.28. \tag{A.23}$$

A.4. 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, which can be calculated according to the following formulas:

$$HDG = TR + \arcsin\left(\frac{WS\cdot\sin(TR-WD)}{IAS}\right), \tag{A.24}$$

$$GS = IAS\cdot\cos(HDG-TR) + WS\cdot\cos(TR-WD). \tag{A.25}$$

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

A.5. Ground return identification and evaluation using a digital elevation model

The ground return velocity is an important parameter for beam pointing-angle calibration as well as motion correction accuracy assessment and refinement. Determination of the ground return velocity requires a reliable ground identification, to determine the number of the range gate intersecting with the ground. Further, an accurate determination of the ground height with respect to the range gate center is important, because the ground presents a hard target for the lidar signal.

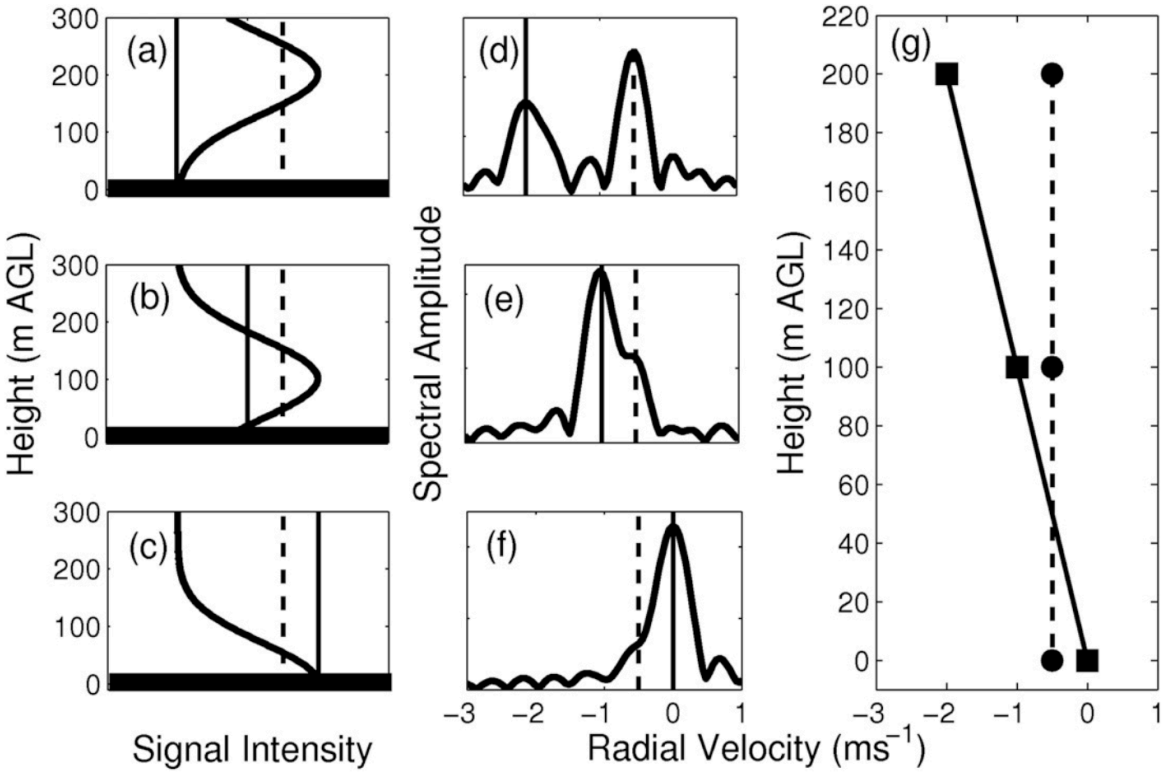

Fig. A.1.: Figure and description from Godwin et al. (2012) illustrating the chirp effect for a similar lidar system: "A schematic of (a)–(c) the variation with height of the intensity of a Gaussian-shaped pulse as it approaches the surface, (d)–(f) the corresponding periodograms that indicate the radial velocities associated with the ground return (solid line) and the aerosol return (vertical dashed line), and (g) the corresponding radial velocity profiles. The radial velocity associated with the aerosol return [dashed line at $-0.5\,\mathrm{m\,s^{-1}}$ in (d)–(g)] remains constant while the radial velocity associated with ground return (solid line) linearly shifts from -2 to $0\,\mathrm{m\,s^{-1}}$ at the ground." © American Meteorological Society. Used with permission.

Due to the hard target return, the measured Doppler velocities suffer from a chirp effect, which systematically biases the measured Doppler velocities depending on the distance of the range gate center to the ground. The chirp effect is caused by a systematic change in the frequency of the laser during the transmission of every pulse. The chirp effect of a similar lidar system is extensively discussed in Godwin et al. (2012), who attempt to extend the retrieval of atmospheric velocities towards the ground. This extension is difficult due to the presence of the ground return signal (present in the lowest $100-200$ m above ground already, due to the Gaussian pulse shape), which overwhelms the aerosol signal due to its substantially higher return reflectance. The first figure from Godwin et al. (2012) is reproduced with permission in Fig. A.1 to illustrate the chirp effect. According to Godwin et al. (2012), the laser frequency decreases systematically from higher to lower frequencies over the duration of one pulse. For hard targets, this systematic change introduces a chirp effect, as the returned frequency (and thereby measured

velocity) depends on the distance of the pulse relative to the ground (for diffuse targets this is not a problem due to the homogeneous distribution of scatterers along the beam). Extending the analysis of Godwin et al. (2012), this work shows that a bias-free ground return velocity can be retrieved once the pulse center has passed the ground, as expected (this was not investigated by Godwin et al. (2012) because they did not utilize the measured ground return velocity but instead tried to suppress its effect). Thereby, the unbiased ground return velocity can be used for motion correction refinement and other analysis.

Due to the chirp effect, the ground return identification requires a reliable estimation of the distance of the range gate center to the ground. The ground determination and velocity estimation implemented in this work uses a two-step procedure in order to obtain unbiased high quality ground return velocities (no extension of the aerosol velocity retrieval towards the ground is attempted). First, a DEM is utilized to identify the number of the range gate intersecting with the ground, similarly to Ellis et al. (2019). Through DEM usage, the distance of the range gate center relative to the ground level is also obtained, thereby the chirp effect can be controlled. Second, the obtained ground return velocities are quality controlled in a similar fashion as in the general quality control procedure. Quality control is especially important, as subsequently the 1 s moving averaged ground return velocities are used to refine the motion correction procedure (Sec. 6.2). Thereby, an erroneous ground return velocity would affect all of the measured lidar velocities along the beam, which needs to be avoided.

The ground elevation at all lidar beam positions H_{TOP} is determined using data from the Shuttle Radar Topography Mission (SRTM, Van Zyl, 2001). The SRTM data has a horizontal resolution of approximately 30 m, a horizontal accuracy of ±8.8 m and a vertical accuracy of ±6.2 m for the region studied (at 90% probability, Rodríguez et al., 2006). Thereby, the resolution is sufficient to allow for high accuracy ground return identification given the lidar range gate length of 69.53 m. It should be noted that pure SNR or backscatter based approaches were tested but do not allow for a sufficiently accurate ground height identification compared to the DEM model based approach.

Using the SRTM ground elevation at the lidar range gate positions (determined in the ECEF coordinate system), a relative distance of every range gate to the ground is calculated based on the height difference between the range gate center and the ground, normalized by the projection of the range gate length onto the vertical ($H^C_{TOP} - H^C(R_{GRD})/(b^C_z \Delta p)$). This relative distance is important, as for example for scans close to the horizontal (during steep turns) the ground range gate will not penetrate the ground for the full range gate length in the vertical due to the shallow incidence. Thereby, an unbiased retrieval will be achieved at a vertical distance less than the full range gate length below the ground.

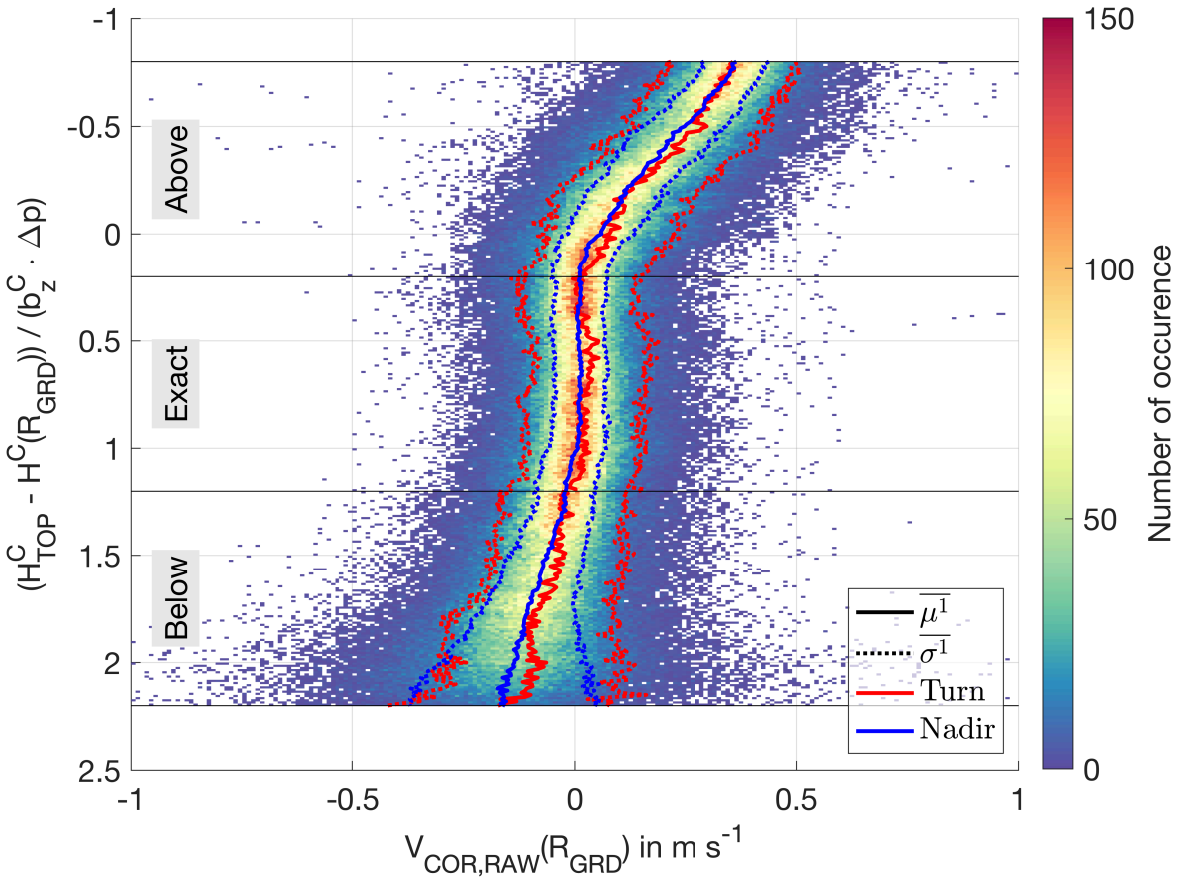

Fig. A.2.: Histogram of measured ground return velocities obtained after raw motion correction. Data is displayed as a function of relative distance from the ground for all nadir and turns legs from all flights. The number of occurrences per Δx =0.01 m s^{-1}, Δy = 0.01 bin is color coded. Additionally shown are the mean (solid line) and standard deviation (dashed lines), separately for nadir and turn legs. Horizontal lines and labeling illustrate the applied sorting scheme into above, exact and below ground range gates. Please note that the chirp effect is of opposite sign compared to Godwin et al. (2012) due to a different convention being used. Extending Godwin et al. (2012), the exact ground range gate velocities show that the radial velocity is unbiased again once the pulse peak has 'passed' the ground (in a temporal sense).

The resulting ground return velocity, as a function of relative distance below the ground, is shown in Fig. A.2. Above the ground, the velocity profile shows the linear bias effect discussed by Godwin et al. (2012). Compared to Godwin et al. (2012) the sign convention is inverted, but also the magnitude of the bias is reduce to values of 0.4 m s^{-1} at a relative distance of 0.8 above the ground. Godwin et al. (2012) use a slightly different lidar system and extend their analysis further up above the ground, explaining the observed difference. Despite the bias, the profile is well centered and the standard deviation is not increased, demonstrating the accuracy of the height determination (small errors in height assignment would have a noticeable effect on the standard deviation due to the linear change with height). Extending Godwin et al. (2012), the figure shows that once the pulse peak passes the ground (in a temporal sense, e.g. the pulse is

Tab. A.1.: Statistics of the quality controlled ground return velocity for all flights including overall availability.

Flight number	Availability in %	Nadir $\overline{\mu^1}$ in $\mathrm{m\,s^{-1}}$	Nadir $\overline{\sigma^1}$ in $\mathrm{m\,s^{-1}}$	Turn $\overline{\mu^1}$ in $\mathrm{m\,s^{-1}}$	Turn $\overline{\sigma^1}$ in $\mathrm{m\,s^{-1}}$
1	95.7	-0.01	0.06	0.04	0.12
2	90.0	0.01	0.05	-0.02	0.12
3	86.2	0.00	0.07	0.01	0.13
4	91.9	0.01	0.06	0.03	0.16

centered on the ground, the retrieval is not literally from below the ground in a spatial sense), the profile remains constant with height and unbiased. Based on the observed profile, the ground range gate is defined to be in the region -0.2 to -1.2 relative height units. Only ground return velocity values from range gates within this region are used for ground return based motion correction refinement. Below this region, again a biased return profile develops, although with smaller magnitude. Further, the number of retrieved velocities starts to decay noticeably and scatter increases (both due to the strongly reduced signal intensity).

Due to the more accurate motion correction during nadir legs (Sec. 6.2.IV), the standard deviation is much lower for nadir legs compared to turns. The standard deviation remains constant with height. The observable scatter of the ground return is due to errors in motion correction, mis-identification of the ground return, radial velocity noise and non-stationary ground targets. A separation and quantification of the importance of each of the drivers is not easily possible. However, the analysis of all measured ground return velocities, as in figure 6.9 as well as figures 6.10, 6.11, 6.15, provides sufficient evidence that the main contribution to ground return velocity is due to errors in motion correction (which is thereby corrected using the ground return) as well as radial velocity noise. Fortunately, the radial velocity noise is low, because the ground return signal strength is generally high, its impact is further reduced for the refined motion correction as the 1 s moving average is used.

After ground return identification, for refined motion correction only (not for the general ground return velocity based quality assessment), the obtained values are further quality controlled in a manner similar to the general quality control procedure. To this end, all values exceeding $\pm 1\,\mathrm{m\,s^{-1}}$ are discarded and only values with a minimum SNR of $-5\,\mathrm{dB}$ are used (in order to restrain the retrieved values for low uncorrelated noise, see Fig. 6.7). Last, a spike filter is applied which removes all spikes with magnitude $> 0.3\,\mathrm{m\,s^{-1}}$ ($> 1\sigma$ of the maximum expected uncorrelated noise level, $> 3\sigma$ of the average expected uncorrelated noise level).

The overall statistics of the available, quality controlled ground return velocities for refined motion correction are given in Tab. A.1. The ground return availability is generally high and but

can be impeded by the presence of clouds. The average ground return is mostly bias-free ($|\overline{\mu^1}| < 0.05\,\mathrm{m\,s^{-1}}$ also for non-quality controlled data), confirming the quality of the beam pointing-angle calibration. As discussed above, the standard deviation of the ground return velocity is much reduced for nadir legs compared to turns.

A.6. Pitch, roll, yaw maneuvers

The usage of PRY calibration maneuvers for lidar data analysis is introduced as a part of this work, based on concepts provided by Lenschow (1986). Haimov and Rodi (2013) utilize similar procedures for ADR beam-pointing angle calibration, for an overview of calibration flight maneuvers to determine nose-boom wind measurement accuracy see Lenschow (1986) and Wendisch and Brenguier (2013)[A1]. A PRY maneuver consists of pilot-induced periodic oscillations around the individual aircraft axis in ARD. First, pitch oscillations are conducted, reaching absolute values of the pitch angle up to 10°. Second, roll oscillations are conducted, reaching absolute values of the roll angle up to 20°. Third, yaw oscillations are conducted, reaching absolute values of the roll angle up to 10°. For an exemplary PRY maneuver and the associated aircraft attitude angles see Fig. 6.9. PRY maneuvers exceed the variations in aircraft attitude angles experienced during standard measurements legs used for meteorological analysis by up to one order of magnitude. Thereby, PRY maneuvers can serve as a lower limit for measurement accuracy during strongly disturbed flight conditions, e.g. due to severe turbulence. Due to the rapid but defined oscillation cycles, systematic errors in the INS reported motion correction data can be detected, as well as its accuracy assessed under extreme conditions.

A.7. Vertical wind variance uncertainty due to horizontal wind contribution

The contamination of the measured vertical wind variance due to non-nadir beam pointing is based on Strauss et al. (2015) in this work. The following formulas for determination of $\overline{\beta^2_{HC}}, \overline{\varepsilon^2_{HC}}$ are taken from their work, the assumptions and derivation are not repeated for the sake of brevity but can be found in their study. The u^A, v^A wind components refer to the along- and across-track wind components, respectively. $\overline{\beta^2_{HC}}, \overline{\varepsilon^2_{HC}}$ are calculated as

$$\overline{\beta^2_{HC}} = \overline{\theta^2}(u^A_{max})^2 + \overline{\theta^1}^2\overline{v_r^2} + \overline{\theta^2}\,\overline{v_r^2} + \overline{\phi^2}(v^A_{max})^2 + \overline{\phi^1}^2\overline{v_r^2} + \overline{\phi^2}\,\overline{v_r^2},$$

$$\overline{\varepsilon^2_{HC}} = 2|\overline{\theta^1}| \cdot 2\sigma_{r_{uw}}\overline{v_r^2} + 2|\overline{\phi^1}| \cdot 2\sigma_{r_{vw}}\overline{v_r^2} + 2|\overline{\theta'\phi'}| \cdot 2\sigma_{r_{uv}}\overline{v_r^2} + 2|\overline{\theta'\phi'}| \cdot u^A_{max}v^A_{max} + 2|\overline{\theta}\,\overline{\phi}| \cdot 2\sigma_{r_{uv}}\overline{v_r^2}.$$

In the above, $\sigma_{r_{uv}}$ is the standard deviation of the correlation coefficient $r_{uv} = \overline{u'^A v'^A}/(\overline{\sigma_u^1 \sigma_v^1})$. This correlation coefficient is unknown at the point of the lidar measurements, therefore it is determined as the average in-situ value from all aircraft measurements inside the boundary layer for all flights (as in Strauss et al., 2015). The values obtained are $\sigma_{r_{uv}} = 0.30$, $\sigma_{r_{uw}} = 0.38$, $\sigma_{r_{vw}} = 0.36$. Further, in the study of Strauss et al. (2015), u^A_{max}, v^A_{max} are the maximum expected along- and across-track horizontal wind components for each flight leg. This rather coarse approximation has to be assumed in their study because they analyze mountain waves and rotors in complex terrain, where the wind profile changes rapidly (and is seldom available in their study). In this study, however, the linearly changing horizontal wind field assumption can be applied due to the less turbulent conditions and the wind profiling conducted before and after each nadir flight leg. Thereby, the horizontal wind contribution to the retrieved radial velocity can be removed in post-processing and consequently the assumption on u^A_{max}, v^A_{max} can be relaxed. Following this, the u^A_{max}, v^A_{max} are determined as the maximum uncertainty associated with the horizontal component of the surrounding wind profiles (using the ADLS developed data-driven uncertainty estimation). The effect of this horizontal wind removal is also discussed in Sec. 6.2.VI, where the results from the reduced uncertainty estimation are compared to those from the original Strauss et al. (2015) estimation.

A.8. Formulas associated with the PVA correction using the modified azimuthal structure function technique

The theory applied for PVA correction and explained in the following is based on Frehlich (1997); Frehlich and Cornman (2002) and Frehlich et al. (2006), however it is modified for application to the vertical wind based on information given in Houbolt et al. (1964) and Lothon et al. (2006). As in Frehlich et al. (2006), the measurement structure function is calculated orthogonal to the beam pointing direction. Thereby, the separations s are given in the direction of the beam movement. Consequently, the structure function is defined as

$$D_v(R, k\Delta s) = \langle [v'(R, s) - v'(R, s + k\Delta s)]^2 \rangle$$

for the radial velocity fluctuations $v'(R, s)$ (deviations from the mean) as a function of range gate R and separation s, where $k\Delta s$ is determined as the separation between measurements along the flight path in the ECEF coordinate system.

A simple model of the structure function for turbulence over a two-dimensional plane is given by Hinze (1959); Frehlich et al. (2006) as

$$D_v(R,s) = 2\overline{w^2}[\Lambda(q/L_0) + \Lambda_D(q/L_0)(1-(R^2/q^2)],$$

with the outer scale of turbulence L_0 and $q = \sqrt{R^2+s^2}$. With the von-Karman model (von Karman, 1948), $\Lambda(x)$ is the longitudinal component given as

$$\Lambda(x) = 1 - \frac{2^{2/3}}{\Gamma(1/3)} x^{1/3} K_{1/3}(x) = 1.0 - 0.592548 x^{1/3} K_{1/3}(x)$$

and $\Lambda_D(x)$ is the transversal component:

$$\Lambda_D(x) = \frac{x^{4/3}}{2^{1/3}\Gamma(1/3)} K_{2/3}(x) = 0.29627426 K_{2/3}(x).$$

Hereby, $\Gamma(x)$ are the Gamma function and $K_{1/3}(x), K_{2/3}(x)$ the modified Bessel functions of order $1/3$ and $2/3$ respectively. The outer scale L_0 is connected to the integral scale $\mathcal{L}$ through

$$\mathcal{L} = \frac{\sqrt{\pi}\Gamma(5/6)}{\Gamma(1/3)} = 0.7468343 L_0.$$

It must be noted (but has not been mentioned in studies using the technique so far) that the above definitions, based on Frehlich et al. (2006), implicitly assumes that the outer scale L_0 is the length scale in the longitudinal direction of the flow. Therefore, the technique needs to be adopted for application to the vertical wind, which is transversal and not longitudinal. The difference between the longitudinal length scale L_u and the transversal length scale L_w is given as $L_u = 2L_w$ for isotropic conditions (Batchelor, 1948; Houbolt et al., 1964).

The assumption of isotropy is not necessarily fulfilled in the atmospheric boundary layer, the degree of violation and its effect on the length scales are evaluated by Lenschow and Stankov (1986). Two reasons warrant usage of a factor 2 in this work. First, any deviations from this ratio are usually unknown at the point of measurement, making this assumption the best choice. Second, any potential violations and their effect are captured by a degraded quality of the raw von-Karman model fit in comparison to the directly obtained length scales estimates. Thereby, any potential deviation from the factor 2 should be captured as uncertainty.

The difference between longitudinal and transversal length scale must be accounted for in the transversal von-Karman model when applying the Frehlich et al. (2006) azimuthal technique to the vertical wind. The difference between the scales and its implications on the von-Karman

model is documented in-depth in Houbolt et al. (1964, Appendix B, p. 64-66), therefore it is not elaborated further here. Not accounting for the factor 2 results in an erroneous PVA correction (systematically increased length scales), which can possibly explain the unexpected PVA correction effect on length scales documented by Lothon et al. (2006, p. 527, although they distinguish between the longitudinal and transversal length scales the difference was maybe not considered in the PVA correction) and Brugger (2014, his Fig. 7.7) when using the transversal model. Also note that the transversal model is applicable for both along-wind and cross-wind directions as the vertical wind is always transversal to the mean flow.

Using the quantities $\overline{w^2}, L_{0,raw}$ (from $\mathcal{L}$ according to Eq. 3.26), calculated from the definition (e.g. without correction of the PVA averaging effect) in the above formula, a raw von-Karman model fit D_{raw} can be constructed for illustrative purpose.

At the same time, according to Frehlich et al. (2006) the PVA weighted model of the structure function is given as

$$\hat{D}_{wgt}(R, s; \overline{w^2}, L_0) = 2\overline{w^2} G(s/\Delta p, \mu, \chi),$$

where the semicolon separates input arguments from model parameters. Hereby, $\mu = \sqrt{2ln(2)}\Delta p/\Delta r, \chi = \Delta p/L_0$ and

$$G(m, \mu, \chi) = 2\int_0^\infty F(x, \mu)[\Lambda(\chi\sqrt{m^2+x^2}) - \Lambda(\chi x) + \Lambda_D(\chi\sqrt{m^2+x^2})m^2/(m^2+x^2)]dx,$$

where the adjusted integration limits and structure function input variables compared to Frehlich et al. (1998) should be noted. F is given by Frehlich et al. (1998) as

$$\begin{aligned} F(x, \mu) = &\frac{1}{2\sqrt{\pi}\mu}\{\exp(-\mu^2(x+1)^2) + \exp(-\mu^2(x-1)^2)\} \\ &+ \frac{x}{2}\{\mathrm{erf}(\mu(x+1)) + \mathrm{erf}(\mu(x-1)) - 2\,\mathrm{erf}(\mu x)\} \\ &- \frac{1}{\sqrt{\pi}\mu}\exp(-\mu^2 x^2) + \mathrm{erf}(\mu(x+1))/2 - \mathrm{erf}(\mu(x-1))/2. \end{aligned}$$

where exp, erf are the standard exponential and error function, respectively.

Assuming homogeneous turbulence and uncorrelated estimation error $E(R)$, an estimate of the azimuth structure function is given by

$$\hat{D}_{wgt}(R, s) = \hat{D}_{noi}(R, s) - E(R).$$

The estimation error $E(R) = 2\overline{\beta_{rv}^2}(R)$ is approximated through extrapolation of the autocovariance function to lag zero using a linear fit through the first four lags (Sec. 6.2.III). The measured structure function is calculated according to

$$\hat{D}_{noi}(R, k\Delta s) = \{V'_{PST,GRD}(R, s) - V'_{PST,GRD}(R, s + k\Delta s)\}^2,$$

with $V'_{PST,GRD}$ as the deviations from the mean.

A.9. Random and systematic error of statistical moments due to finite measurement lengths

Both systematic and random errors are introduced if statistical quantities of the vertical wind are calculated from measurements of finite length, because the desired ensemble average has to be replaced by the temporal mean (Lenschow and Kristensen, 1984; Lenschow and Stankov, 1986; Lenschow et al., 1994). Hence, the associated uncertainties need to be estimated. First, the ensemble average of the time mean is not equal to the ensemble average, $\langle\overline{\mu^2}(T)\rangle \neq \overline{\mu^2}$. The systematic uncertainty due to this estimation is given by Lenschow et al. (1994) as

$$\frac{\langle\overline{\mu^2}(T)\rangle}{\overline{\mu^2}} \approx 1 - 2\frac{\mathcal{T}}{T}. \tag{A.26}$$

The ratio $\frac{\mathcal{T}}{T}$ expresses the sample size of independent data which can be gathered (further related to the number of turbulent eddies which are averaged over). The relationship in Eq. A.26 shows that there is a systematic underestimation of the second-order moment if measured using time series of limited extent. Fortunately, this systematic offset can be estimated and becomes negligible for sufficiently long measurement duration compared to the integral scale.
Second, another important source of uncertainty is the random error when estimating the ensemble mean from the second-order central moment of $w(t)$. The individual realizations $\overline{\mu^2}(T)$ are distributed around $\langle\overline{\mu^2}(T)\rangle$ with a random variance

$$\overline{\varepsilon_{ran}^2}^2(T) = \langle[\overline{\mu^2}(T) - \langle\overline{\mu^2}(T)\rangle]^2\rangle. \tag{A.27}$$

According to Lenschow and Kristensen (1984); Lenschow et al. (1994), the random variance can be estimated as

$$\overline{\varepsilon_{ran}^2}^2(T) = 4\frac{\overline{\mu^2}^2}{T}\int_0^\infty \rho^2(\tau)d\tau, \tag{A.28}$$

which simplifies to

$$\overline{\varepsilon_{ran}^2}(T) = \sqrt{2\overline{\mu^2}^2 \frac{\mathcal{T}}{T}}, \tag{A.29}$$

if an exponential autocorrelation function (giving $\int_0^\infty \rho^n(\tau)d\tau = \frac{\mathcal{T}}{n}$) is assumed. Please note that for the vertical wind $\overline{\mu^2}$ has units of $\mathrm{m^2\,s^{-2}}$ and accordingly $\overline{\varepsilon_{ran}^2}(T)$ has units of $\mathrm{m^2\,s^{-2}}$. Again, longer averaging times reduces the uncertainty magnitude (because more eddies are averaged, as long as the ergodicity assumption holds) and also enables a more reliable uncertainty estimation as both the integral length scale and $\overline{\mu^2}$ need to be estimated. The random uncertainty is much larger than the systematic uncertainty for small values of $\frac{\mathcal{T}}{T}$ (the usual case, as averaging over many times the integral scale is necessary for reliable measurements), due to the square root present when evaluating for $\overline{\varepsilon_{ran}^2}(T)$ (discussed by Lenschow et al., 1994, in-depth also for disjunct sampling). The above framework, adopted for the first order moment, is also used to estimate the random uncertainty associated with the measurement of the mean, as is done for example by Lenschow and Kristensen (1984), Lenschow and Stankov (1986) and Lenschow et al. (1999):

$$\overline{\varepsilon^1}_{ran}(T) = \sqrt{2\overline{\mu^2}\frac{\mathcal{T}}{T}}. \tag{A.30}$$

Please note that in the measurements the true ensemble means $\overline{\mu^1}, \overline{\mu^2}$ are always substituted by $\overline{\mu^1}(T), \overline{\mu^2}(T)$ due to the impossibility of estimating $\overline{\mu^1}, \overline{\mu^2}$ directly.

B. Additional tables

Tab. B.1.: Overview of used acronyms and their meaning. For parameters which are varied the standard values used for system setup (STP) and wind profile retrieval strategy (RET) are marked in red.

Simulator settings			
Name	Description	LES set A	LES set B
AAL	Aircraft flight altitude	1100 m	
AVG	Along-track averaging distance	325, 650, 975, 1300, 2600 m	
BLA	Sector blanking	0°, 90°, 135°, 180°, 225°, 270°	
ELE	Scan elevation angle	30°, 40°, 50°, 60°, 65°, 70°, 75°, 80°	
FME	Lidar measurement frequency	1, 5, 10 Hz	
HGT	Height used for retrieval	<800 m	
IAS	Indicated air speed	65 $m\,s^{-1}$	
LEV	Vertical averaging distance	60, 120, 240 m	
ROT	Scan rotation speed	18, 22.5, 30, 45, 90° s^{-1}	
WND	Background wind case	0, 2, 5, 10, 15 $m\,s^{-1}$	5, 10, 15 $m\,s^{-1}$
	Wind field during sampling	frozen	varying
	LES time steps sampled	7	28
	Transects flown per time step	8	3
	Profiles retrieved per transect	6	1
	Time separation between transects	20 minutes	1 minute

Tab. B.2.: List of analyzed flight legs for flight 1 (10 July 2019, flight time 11:35-13:30 UTC). Average quantities given are measured by the aircraft at flight altitude. Gaps between analyzed legs are due to laser calibration procedures, presence of clouds or laser stability issues. An illustration of the flight and the geographic location of the legs is provided in Fig. C.15, an overview of the ADL measured data is given in Fig. C.19.

Leg number	Leg type	Begin time UTC	End time UTC	z_i in m	Min. Alt. in m	Max. Alt. in m	Avg. ϕ in deg	Avg. ψ in deg	Avg. V_m in $\mathrm{m\,s^{-1}}$	Avg. V_α in deg
1	Turn asc.	11:44:45	11:54:00	1000	739	2952	-17.9	177	5.4	311
2	Turn	11:54:00	11:57:20	1000	2952	2975	-16.1	156	7.0	328
3	Nadir	12:00:30	12:07:30	1000	2964	2970	-0.2	349	7.3	328
4	Turn	12:07:40	12:12:30	1000	2969	2990	18.4	174	7.3	329
5	Turn	12:13:50	12:18:10	1000	2973	2996	-28.4	178	7.6	336
6	Turn	12:18:30	12:20:10	1000	2964	2992	-42.2	182	8.0	337
7	Turn	12:20:20	12:20:50	1000	2879	2976	-57.2	205	7.5	342
8	Turn desc.	12:22:00	12:30:00	1000	277	2805	-23.8	169	5.6	328
9	Nadir asc.	12:30:00	12:32:00	1000	279	945	2.9	167	4.8	319
10	Nadir	12:32:00	12:42:00	1000	922	949	-0.3	170	5.3	315
11	Turn	12:47:30	12:50:00	1000	2357	2384	29.5	181	6.5	323
12	Turn	12:54:00	12:56:05	1000	2383	2418	28.6	128	5.6	321
13	Turn	12:59:10	13:00:20	1000	2357	2370	-30.3	201	5.4	324
14	PRY	13:00:30	13:03:45	1000	2359	2441	-0.1	145	5.4	332
15	Other	13:05:45	13:12:45	1000	2360	2374	-6.7	234	4.6	328
16	Turn	13:13:00	13:14:30	1000	2355	2377	43.7	154	4.6	331
17	Nadir desc.	13:15:45	13:25:00	1000	297	2206	-0.8	267	3.0	292

Tab. B.3.: List of analyzed flight legs for flight 2 (18 July 2019, flight time 06:15-08:58 UTC). Average quantities given are measured by the aircraft at flight altitude. Gaps between analyzed legs are due to laser calibration procedures, presence of clouds or laser stability issues. An illustration of the flight and the geographic location of the legs is provided in Fig. C.16, an overview of the ADL measured data is given in Fig. C.20.

Leg number	Leg type	Begin time UTC	End time UTC	z_i in m	Min. Alt. in m	Max. Alt. in m	Avg. ϕ in deg	Avg. ψ in deg	Avg. V_m in $\mathrm{m\,s^{-1}}$	Avg. V_α in deg
1	Other	06:15:00	06:32:00	200	57	1729	0.7	182	NaN	NaN
2	Turn	06:32:20	06:33:10	200	1714	1735	-37.7	218	3.0	151
3	PRY	06:33:20	06:35:20	200	1699	1743	-0.7	180	3.6	136
4	Nadir	06:35:30	06:37:30	200	1719	1744	-0.3	178	3.3	123
5	Turn asc.	06:38:00	06:42:00	500	1830	2917	-17.8	163	5.1	155
6	Nadir	06:43:00	06:49:45	1000	2921	2950	-0.5	203	6.7	170
7	Turn	06:49:50	06:50:40	500	2930	2952	-38.1	143	5.4	169
8	Nadir	06:51:00	06:55:00	200	2925	2953	-0.5	203	5.8	171
9	Turn	06:55:30	06:56:40	200	2921	2943	-37.9	147	5.8	181
10	Nadir	06:59:10	07:05:25	200	2929	2940	-0.4	322	5.5	168
11	Turn	07:05:40	07:06:25	200	2919	2942	-38.6	153	5.7	136
12	Nadir	07:09:30	07:14:45	200	2944	2955	-0.7	325	6.3	132
13	Turn	07:14:55	07:16:20	200	2918	2960	-38.1	194	6.9	124
14	Nadir	07:18:35	07:24:35	200	2935	2948	-0.2	119	6.3	146
15	Turn	07:24:50	07:25:35	200	2916	2945	-39.7	218	6.4	159
16	Nadir	07:25:50	07:34:30	1000	2932	2958	-0.5	122	6.9	159
17	Turn	07:34:45	07:35:45	1500	2920	2949	-40.0	174	7.7	179
18	Nadir	07:36:00	07:40:20	1000	2924	2951	-0.3	26	7.2	183

Table B.3 continued from previous page

19	Turn	07:40:30	07:41:30	200	2916	2936	-40.5	189	6.3	180
20	PRY	07:41:55	07:43:50	200	2931	2972	-1.4	203	6.6	170
21	Nadir	07:43:50	08:01:40	1000	2938	2964	-0.5	203	5.8	185
22	Turn	08:02:00	08:02:40	1000	2932	2956	-38.1	115	4.7	273
23	Turn desc.	08:03:00	08:10:10	1000	863	2914	2.6	193	4.5	238
24	Nadir asc.	08:10:30	08:11:30	1200	910	1163	-0.5	30	1.1	237
25	Nadir	08:11:40	08:15:55	1300	1130	1166	-0.4	30	2.2	217
26	Nadir asc.	08:16:00	08:17:25	1000	1147	1553	0.1	30	3.2	174
27	Nadir	08:17:25	08:19:10	1300	1541	1567	-0.5	34	4.0	167
28	Nadir asc.	08:19:10	08:20:00	1300	1538	1710	-0.5	29	4.5	164
29	Nadir	08:20:00	08:26:10	1600	1620	1708	-0.8	30	5.2	152
30	Nadir	08:27:50	08:31:20	1000	1639	1651	-0.1	22	7.4	135
31	Turn asc.	08:31:30	08:35:20	1000	1682	2865	-13.4	136	9.2	153
32	Turn	08:36:00	08:37:00	1000	2931	2945	-37.5	176	8.8	163
33	PRY	08:37:00	08:38:30	1000	2943	2965	-0.6	164	8.9	161
34	PRY	08:38:50	08:40:15	1000	2968	3003	0.5	253	9.0	156
35	Turn	08:40:25	08:41:10	1000	2955	2986	18.6	276	8.5	155
36	Nadir	08:41:15	08:43:50	1000	2931	2955	-0.3	21	8.5	162
37	Turn desc.	08:44:00	08:49:40	1000	981	2927	-24.7	187	6.9	148
38	Nadir desc.	08:49:40	08:53:45	1000	292	981	-2.1	338	3.5	157

Tab. B.4.: List of analyzed flight legs for flight 3 (18 July 2019, flight time 12:30-15:14 UTC). Average quantities given are measured by the aircraft at flight altitude. Gaps between analyzed legs are due to laser calibration procedures, presence of clouds or laser stability issues (from 12:30:00-13:00:00 UTC the laser temperature was too high and the laser did not work). From 14:46:00-14:47:45 UTC the aircraft data acquisition failed. An illustration of the flight and the geographic location of the legs is provided in Fig. C.17, an overview of the ADL measured data is given in Fig. C.21.

Leg number	Leg type	Begin time UTC	End time UTC	z_i in m	Min. Alt. in m	Max. Alt. in m	Avg. ϕ in deg	Avg. ψ in deg	Avg. V_m in m s^{-1}	Avg. V_α in deg
1	Other	12:30:00	13:00:00	2200	53	1895	-7.1	142	NaN	NaN
2	Turn	13:00:00	13:01:30	2200	1848	1877	-37.5	198	4.1	164
3	PRY	13:01:45	13:03:30	2200	1842	1881	-1.1	203	4.6	161
4	Nadir	13:04:30	13:08:00	2200	1861	1879	-0.2	206	2.3	148
5	Turn	13:08:20	13:09:30	2200	1860	1885	-32.0	217	1.9	169
6	Nadir	13:09:30	13:18:45	2200	1860	1899	-0.5	207	3.4	210
7	Turn	13:19:00	13:19:45	2200	1870	1891	-36.6	134	5.6	264
8	Nadir	13:20:00	13:26:15	2200	1861	1888	-0.5	213	4.8	250
9	Turn	13:26:30	13:27:20	2200	1860	1877	-36.8	139	4.6	261
10	Turn desc.	13:27:30	13:30:20	2200	796	1854	21.0	236	5.0	250
11	Turn asc.	13:30:30	13:36:10	2200	737	2294	13.0	204	5.7	247
12	Turn asc.	13:37:15	13:38:45	2200	2599	2973	14.3	250	10.8	269
13	Nadir	13:39:00	13:44:10	2200	3019	3049	-0.6	19	12.3	276
14	Turn	13:44:20	13:46:10	2200	3008	3054	0.6	177	12.6	275
15	Nadir	13:46:20	13:52:10	2200	3033	3060	-0.2	17	12.5	269
16	Turn desc.	13:52:20	13:54:45	2200	1893	3025	17.3	251	6.6	223

Table B.4 continued from previous page

17	PRY	13:55:00	13:56:35	2200	1878	1916	-1.3	37	6.3	219
18	Nadir	13:57:30	14:07:00	2200	2019	2059	-0.7	209	5.5	230
19	Turn	14:07:15	14:08:30	2200	1995	2045	-39.1	135	8.2	265
20	Nadir	14:08:40	14:13:00	2200	1994	2060	-0.3	21	6.2	276
21	Turn	14:13:00	14:14:30	2200	1987	2039	-30.7	137	3.4	257
22	Nadir	14:14:45	14:20:50	2200	2016	2044	-0.3	293	4.3	264
23	Turn	14:21:00	14:21:55	2200	2016	2035	-37.8	166	4.1	273
24	Nadir	14:22:10	14:28:00	2200	1708	2014	-0.7	293	3.0	265
25	Turn	14:28:20	14:29:20	2200	1655	1709	-38.9	182	2.6	282
26	Nadir	14:32:00	14:35:40	2200	1711	1880	-0.4	121	4.8	285
27	Turn	14:35:45	14:36:45	2200	1866	1887	-36.9	181	5.1	275
28	Nadir	14:36:45	14:43:30	2200	1858	1883	-0.6	119	5.0	274
29	Turn	14:43:45	14:44:45	2200	1858	1880	35.8	185	5.5	270
30	Nadir	14:46:00	14:47:45	2200	1863	1884	NaN	NaN	NaN	NaN
31	Turn	14:48:00	14:49:15	2200	1824	1878	-37.7	142	6.0	274
32	Nadir	14:50:15	14:57:15	2200	2017	2035	-1.0	17	6.9	265
33	Turn	14:57:25	14:58:30	2200	2022	2058	-34.9	170	6.4	227
34	Turn asc.	14:59:45	15:01:45	2200	2457	2949	-25.5	142	5.2	245
35	Turn	15:01:45	15:03:10	2200	2940	2976	-38.2	159	8.1	267
36	PRY	15:03:20	15:04:25	2200	2958	3027	-0.4	301	9.5	266
37	Nadir desc.	15:05:00	15:14:00	2200	246	2801	-2.4	350	6.2	223

Tab. B.5.: List of analyzed flight legs for flight 4 (19 July 2019, flight time 09:40-11:38 UTC). Average quantities given are measured by the aircraft at flight altitude. Gaps between analyzed legs are due to laser calibration procedures, presence of clouds or laser stability issues. An illustration of the flight and the geographic location of the legs is provided in Fig. C.18, an overview of the ADL measured data is given in Fig. C.22.

Leg number	Leg type	Begin time UTC	End time UTC	z_i in m	Min. Alt. in m	Max. Alt. in m	Avg. ϕ in deg	Avg. ψ in deg	Avg. V_m in $\mathrm{m\,s^{-1}}$	Avg. V_α in deg
1	Nadir	09:44:45	09:50:00	1300	757	831	-0.6	47	6.2	256
2	Nadir asc.	09:50:15	09:53:15	1300	874	1903	0.2	50	8.1	257
3	Turn asc.	09:55:00	09:57:30	1300	2495	2951	-17.6	156	8.9	249
4	Nadir	09:58:00	10:07:00	1300	2938	2987	-0.8	69	8.8	239
5	Turn	10:07:15	10:08:30	1300	2949	2981	-38.1	170	8.9	237
6	Nadir	10:08:45	10:16:00	1300	2967	2990	-0.6	337	7.8	239
7	Turn	10:16:10	10:17:15	1300	2952	2978	-39.0	189	6.6	235
8	Turn desc.	10:17:30	10:22:15	1300	1186	2827	-21.1	185	4.6	267
9	Nadir	10:22:30	10:32:45	1300	1108	1148	-0.6	162	4.1	271
10	Turn	10:33:50	10:34:45	1300	1089	1125	-35.8	149	4.0	247
11	Nadir	10:35:00	10:44:45	1300	1077	1110	-0.7	250	4.3	287
12	Turn	10:45:00	10:46:15	1300	1093	1109	35.6	206	3.9	278
13	Nadir	10:46:30	10:55:15	1300	1093	1131	-0.8	339	4.5	261
14	Turn	10:55:30	10:56:30	1300	1087	1120	-37.6	171	5.0	231
15	Nadir desc.	10:56:30	10:58:30	1300	289	1115	9.7	157	5.5	253
16	Other	10:58:30	11:07:45	1300	240	305	-0.9	162	5.4	258
17	Other	11:08:00	11:15:30	1300	368	2897	0.8	272	8.6	251
18	Turn	11:15:45	11:16:45	1300	2918	2946	-40.0	186	7.9	227

Table B.5 continued from previous page

19	Other	11:16:45	11:20:00	1300	1190	2941	-4.6	175	7.6	260
20	Other	11:20:00	11:37:10	1300	343	1190	-2.6	256	6.1	257

C. Additional figures

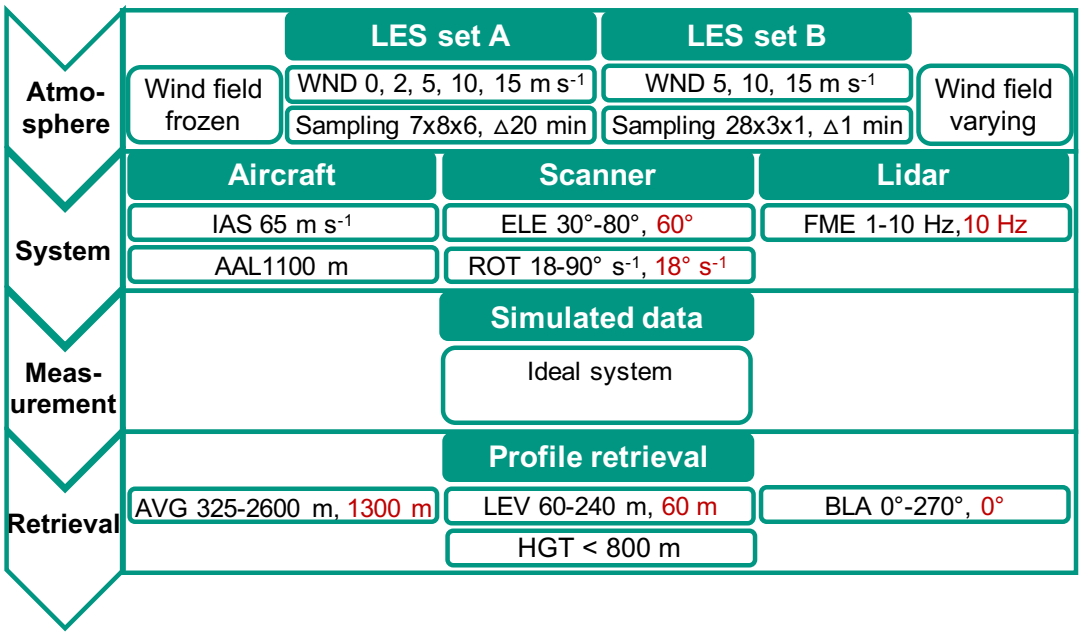

Fig. C.1.: ADLS operation scheme and settings used for the AVAD wind profiling quality analysis. The corresponding abbreviations can be found in Tab. B.1. For parameters which are varied the standard settings for system setup (STP) and wind profile retrieval strategy (RET) are marked in red.

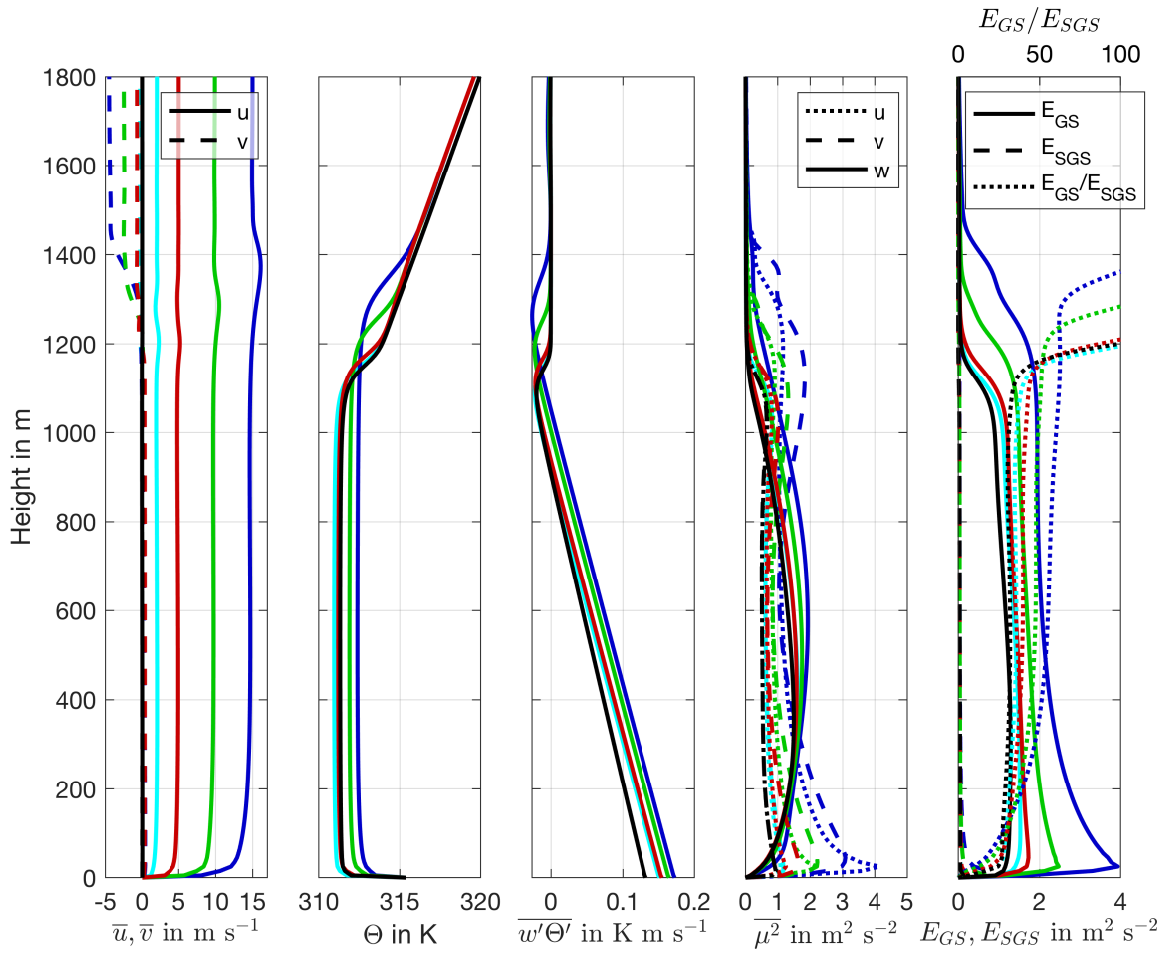

Fig. C.2.: Vertical profiles of the average wind speed, potential temperature, kinematic sensible heat flux, the component-wise wind variance and the grid vs. sub-grid scale turbulent kinetic energy for LES set A. The different background wind cases are color-coded. The corresponding figures for LES set B are provided in Stawiarski (2014), Fig. 5.1 and 5.2.

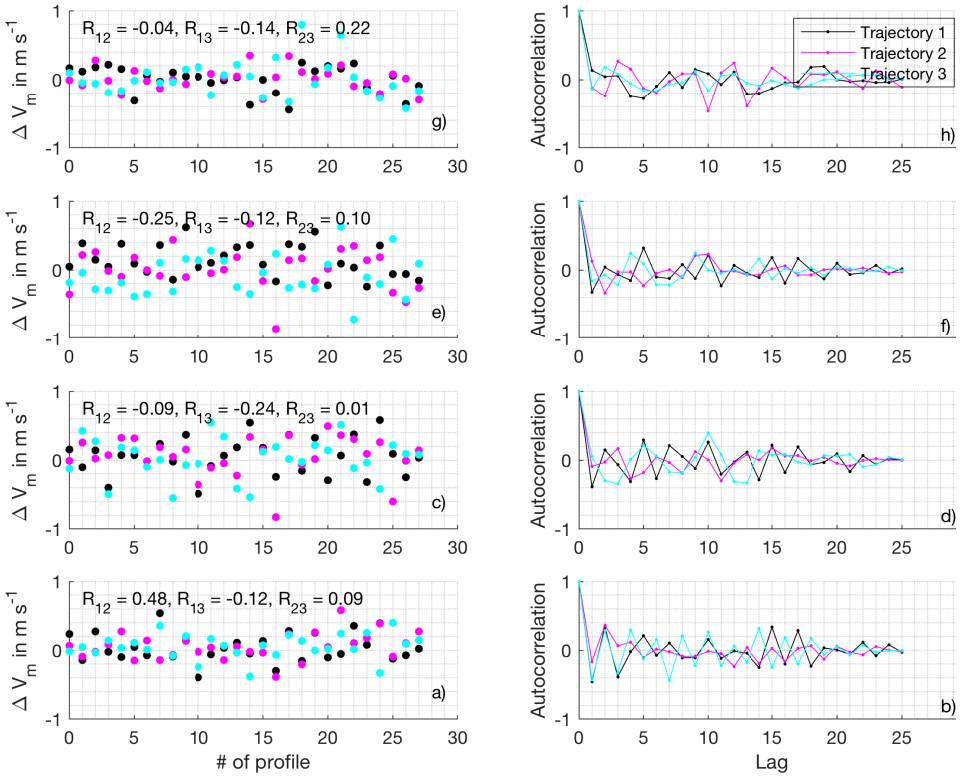

Fig. C.3.: Error correlation analysis for the $5\,\mathrm{m\,s^{-1}}$ background wind case. Panels a), c), e), g) show the retrieval errors for every time step at $60, 320, 500, 760$ m altitude respectively. Color-coded is the number of the parallel trajectory. Additionally given as numbers is the correlation between neighboring trajectories. Panels b), d), f), h) show the autocorrelation of the retrieval errors for each transect.

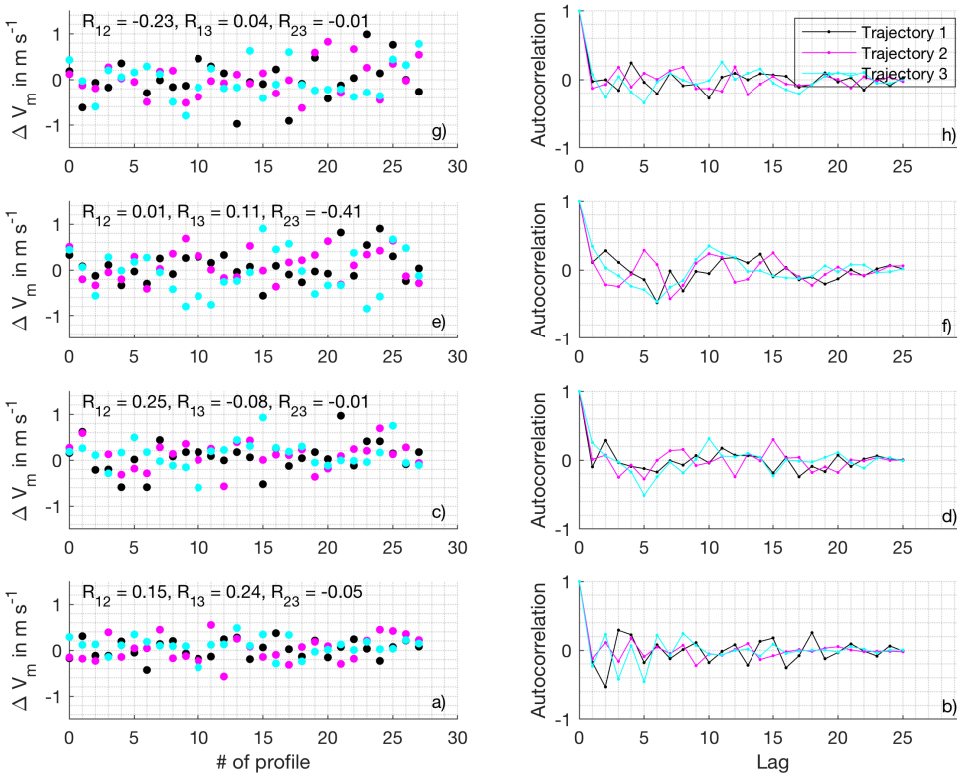

Fig. C.4.: Same as Fig. C.3 but for the 10 m s^{-1} background wind case.

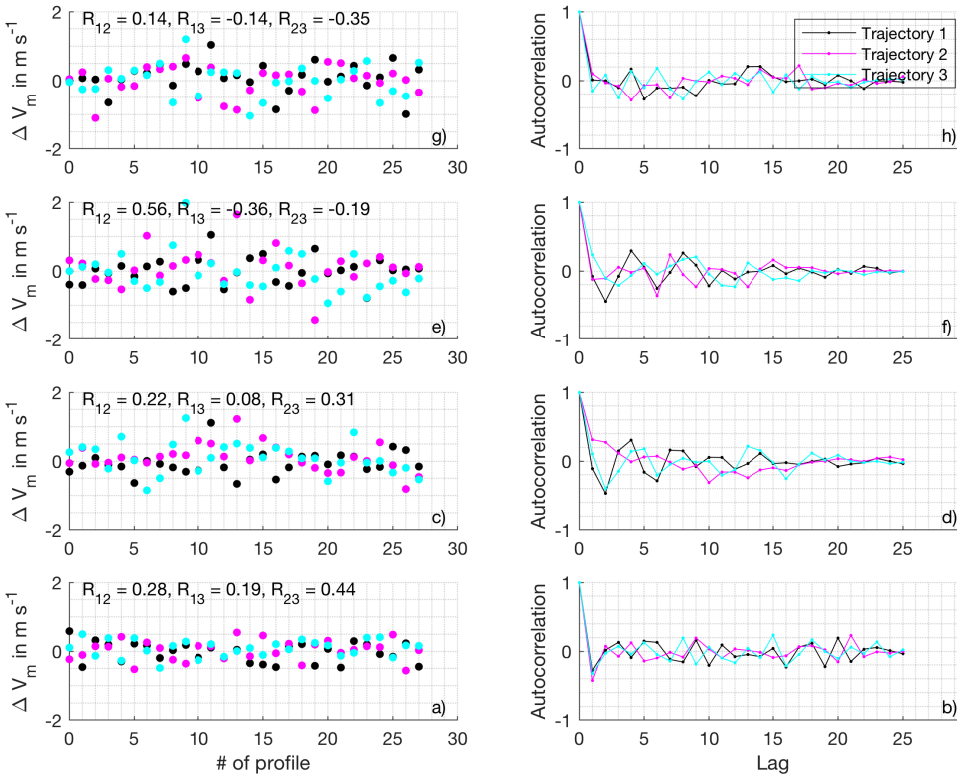

Fig. C.5.: Same as Fig. C.3 but for the 15 m s^{-1} background wind case.

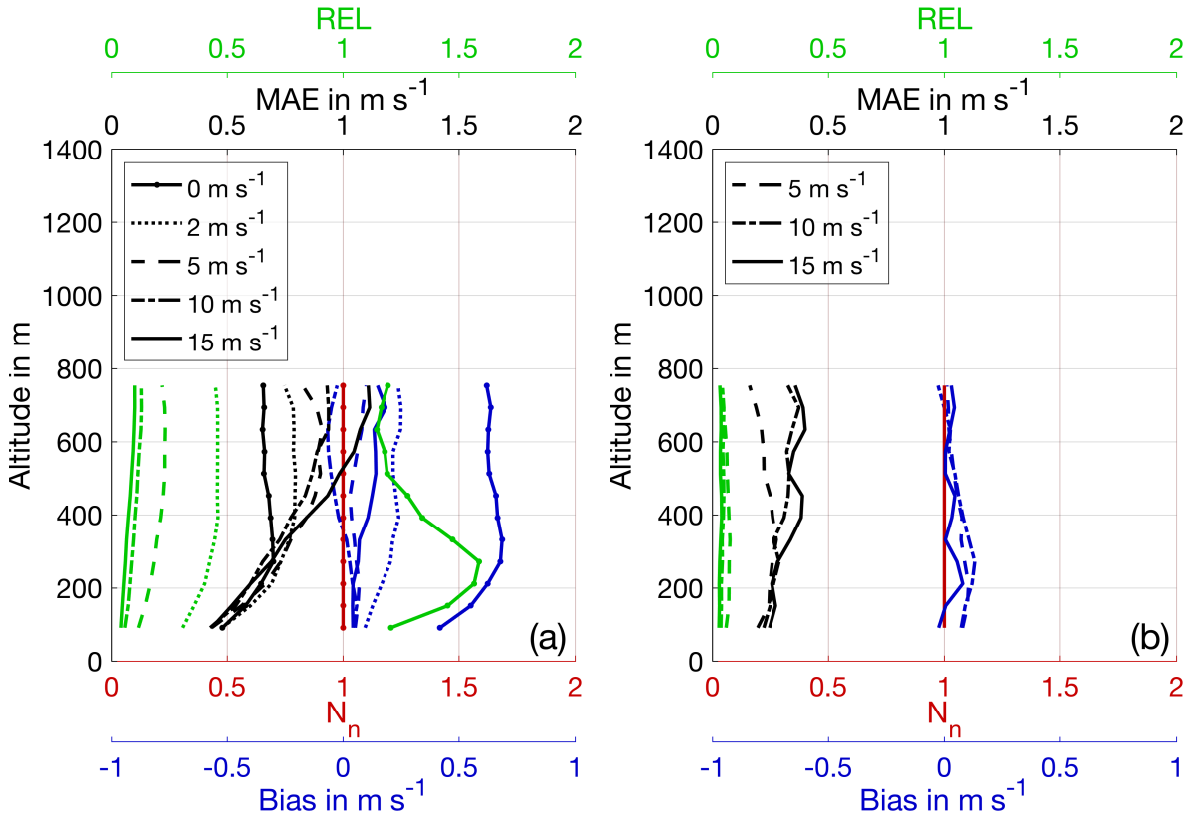

Fig. C.6.: Profile of wind speed retrieval quality parameters MAE, REL, bias and normalized number of retrieved wind profile points N_n as a function of height. Displayed data are for an ideal system using the standard system setup and retrieval strategy for (a) LES set A and (b) LES set B.

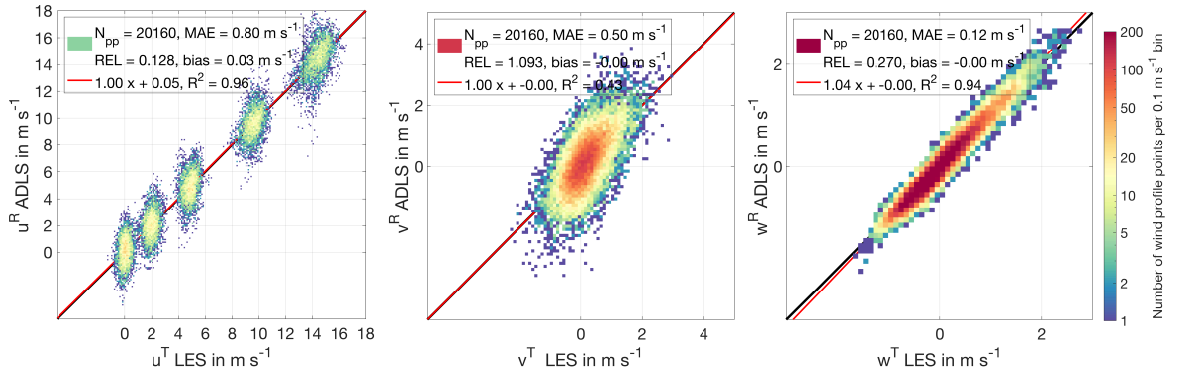

Fig. C.7.: Same as 5.1 but for the individual u, v, w components.

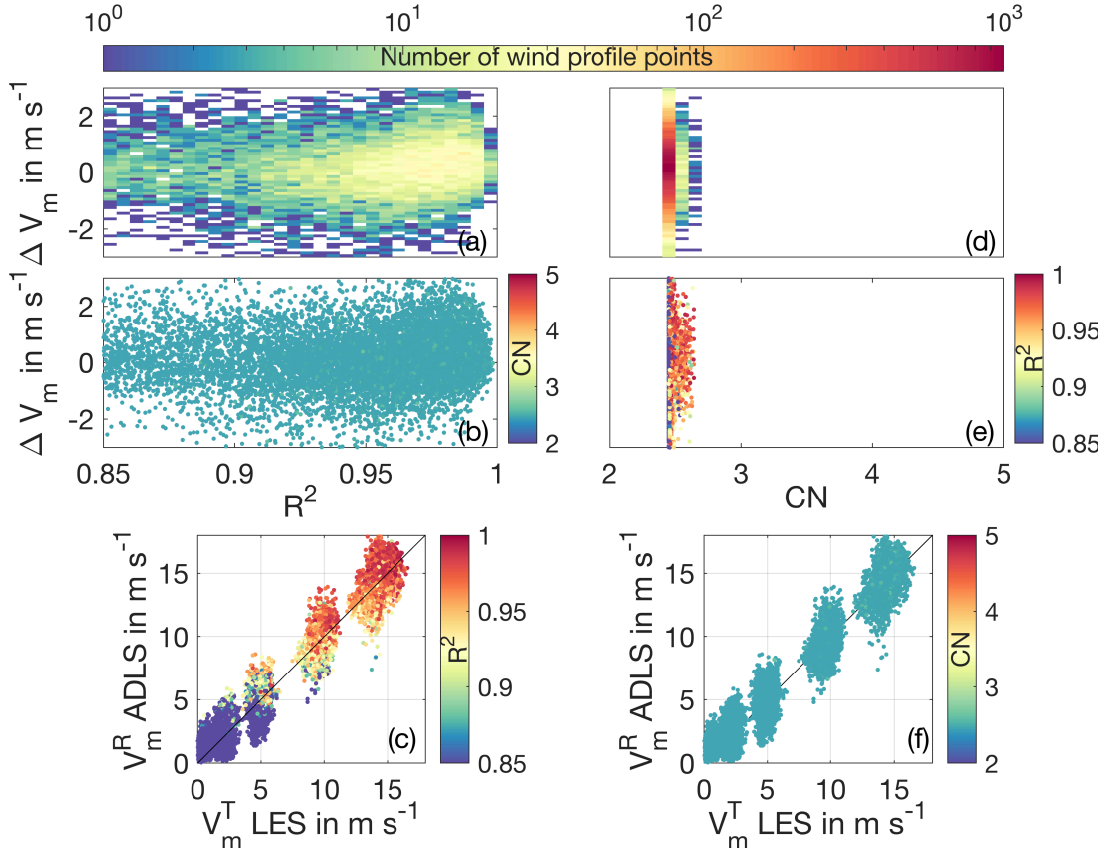

Fig. C.8.: Quality control criteria for the standard system setup and retrieval strategy for LES set A, as in Fig. 5.3 but with adjusted color scale.

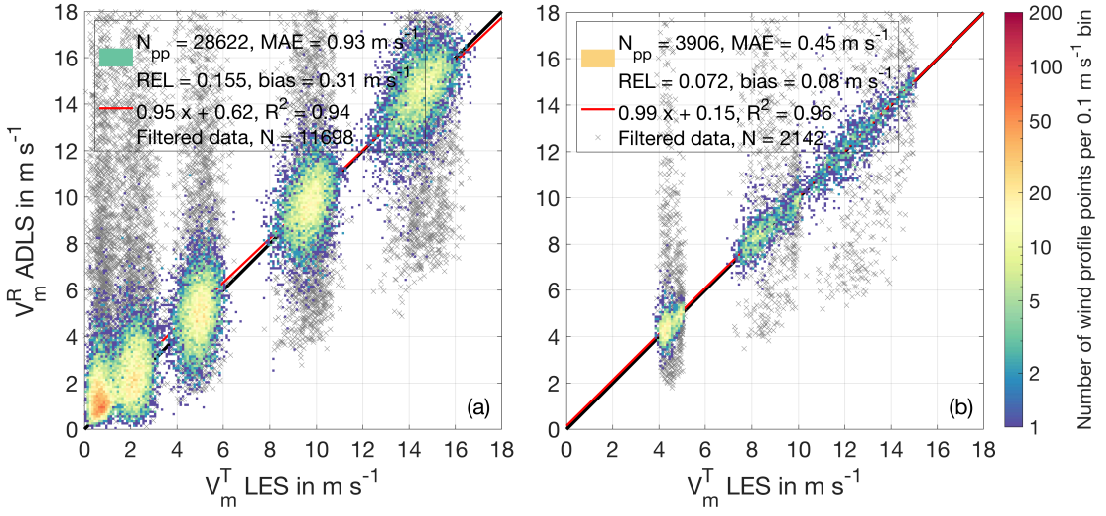

Fig. C.9.: Comparison of LES truth and AVAD retrieved wind speed for an ideal measurement system for (a) LES set A and (b) LES set B using the increased along-track resolution retrieval strategy (AVG = 650 m). Color-coded is the histogram distribution of all measurements. The observable deviations from the 1:1 line reveal the AVAD retrieval error due to the violation of the flow homogeneity assumption. The system setup and retrieval strategy are the same between LES set A and B, more wind profiles can be retrieved from LES set A due to the larger simulation domain.

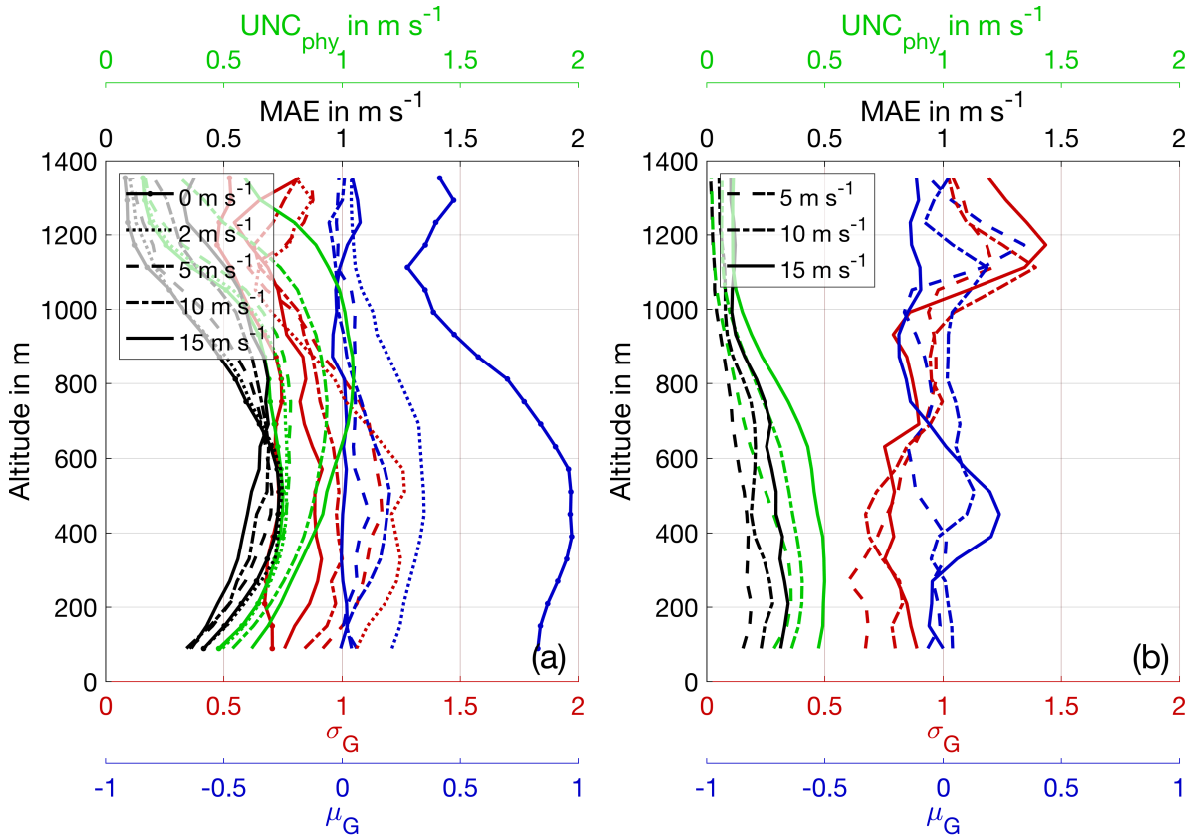

Fig. C.10.: Profile of uncertainty forecast parameters UNC_{phy}, MAE, μ_G and σ_G as a function of height. Displayed data are for (a) LES set A and (b) LES set B and an ideal system with the aircraft flying at an altitude of 1700 m, using the standard system setup and retrieval strategy otherwise.

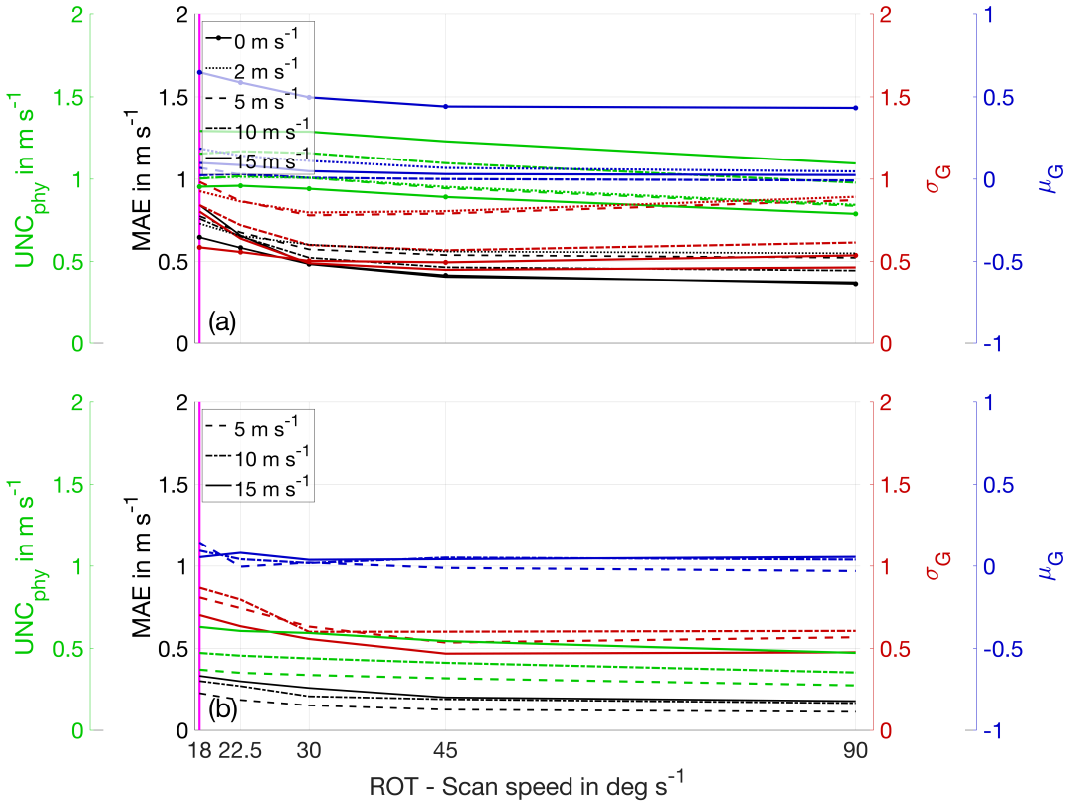

Fig. C.11.: Same as Fig. 5.15 but for five different scan speeds, using STP and RET otherwise.

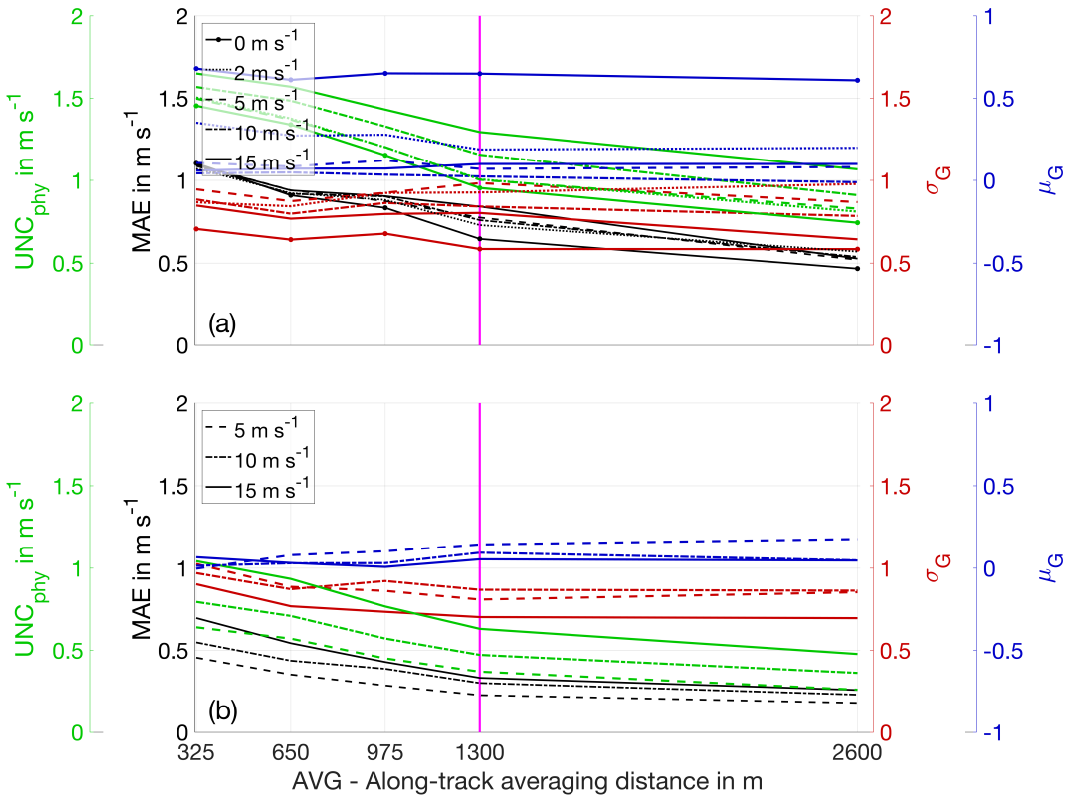

Fig. C.12.: Same as Fig. 5.15 but for three different along-track averaging distances, using STP and RET otherwise.

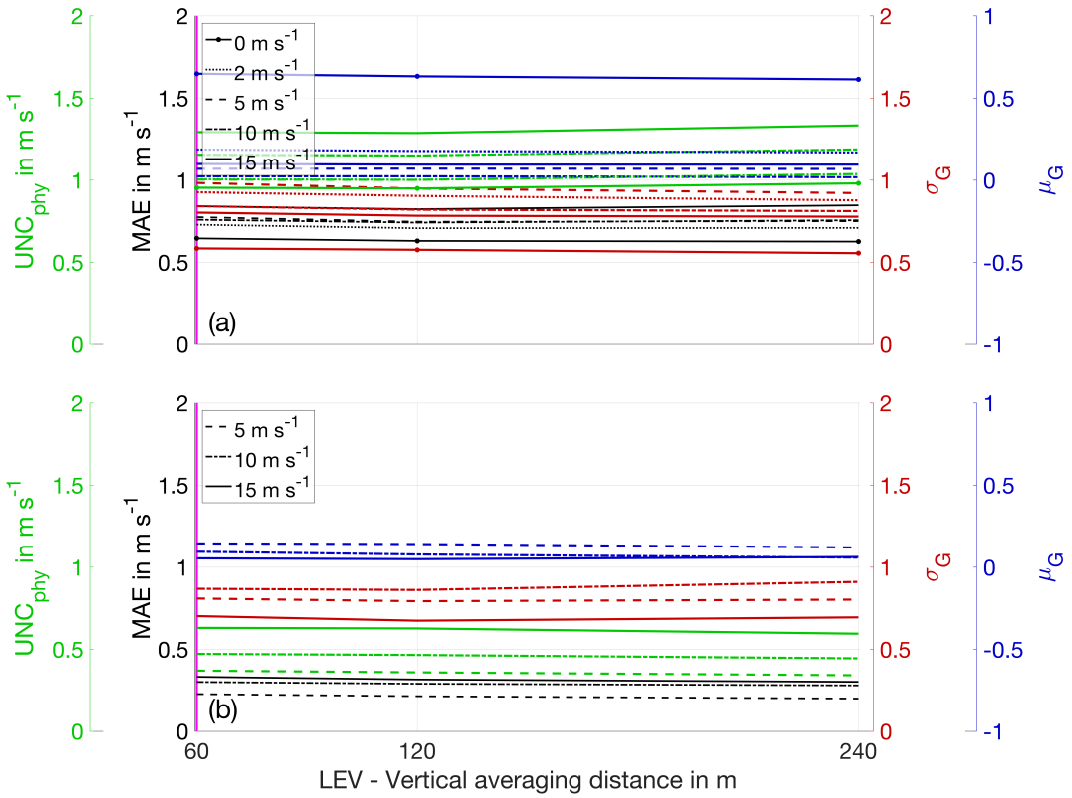

Fig. C.13.: Same as Fig. 5.15 but for three different vertical averaging distances, using STP and RET otherwise.

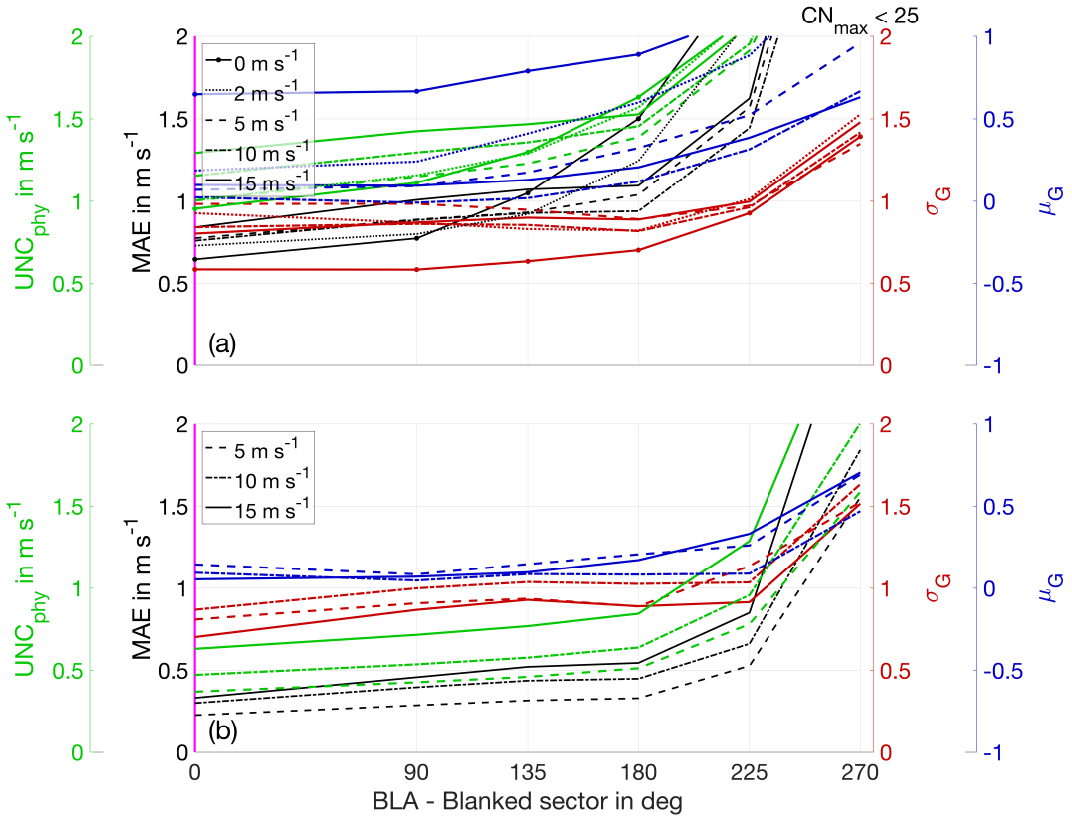

Fig. C.14.: Same as Fig. 5.15 but for four different sector blankings, using STP and RET otherwise.

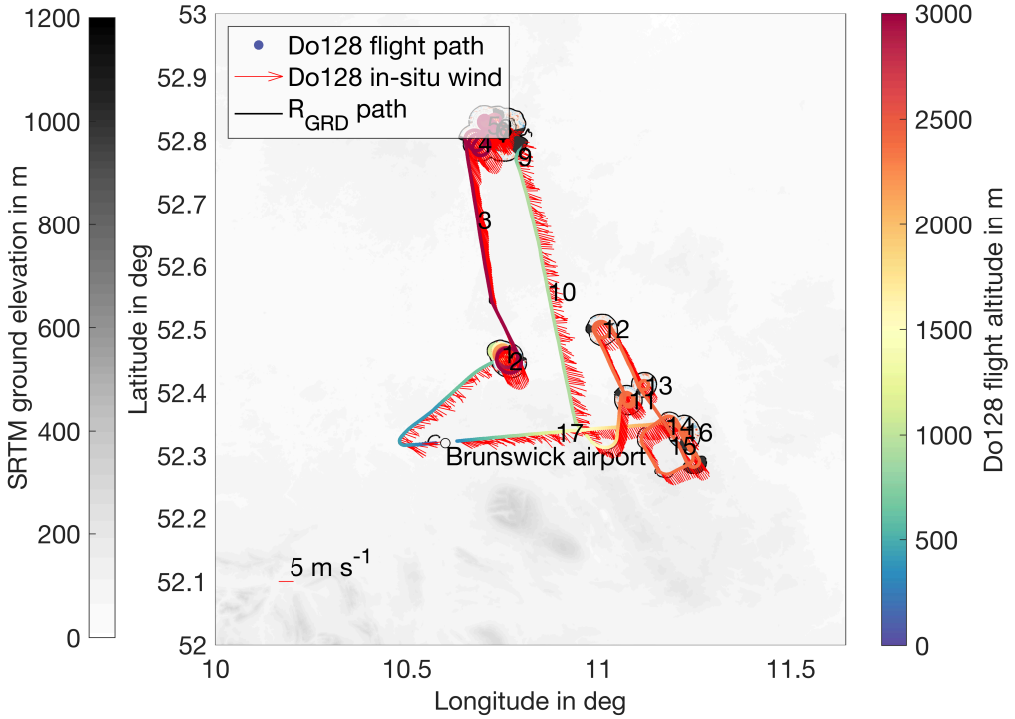

Fig. C.15.: Illustration of flight path for flight 1 (10 July 2019, flight time 11:35-13:30 UTC). The Do128 flight trajectory is shown with the aircraft altitude color-coded. Quivers show the in-situ measured wind speed and direction at flight altitude with a 5 s average applied. Additionally shown is the path traced by the lidar ground range gate. Numbers correspond to the flight legs detailed in Tab. B.2, an overview of the ADL measured data is given in Fig. C.19.

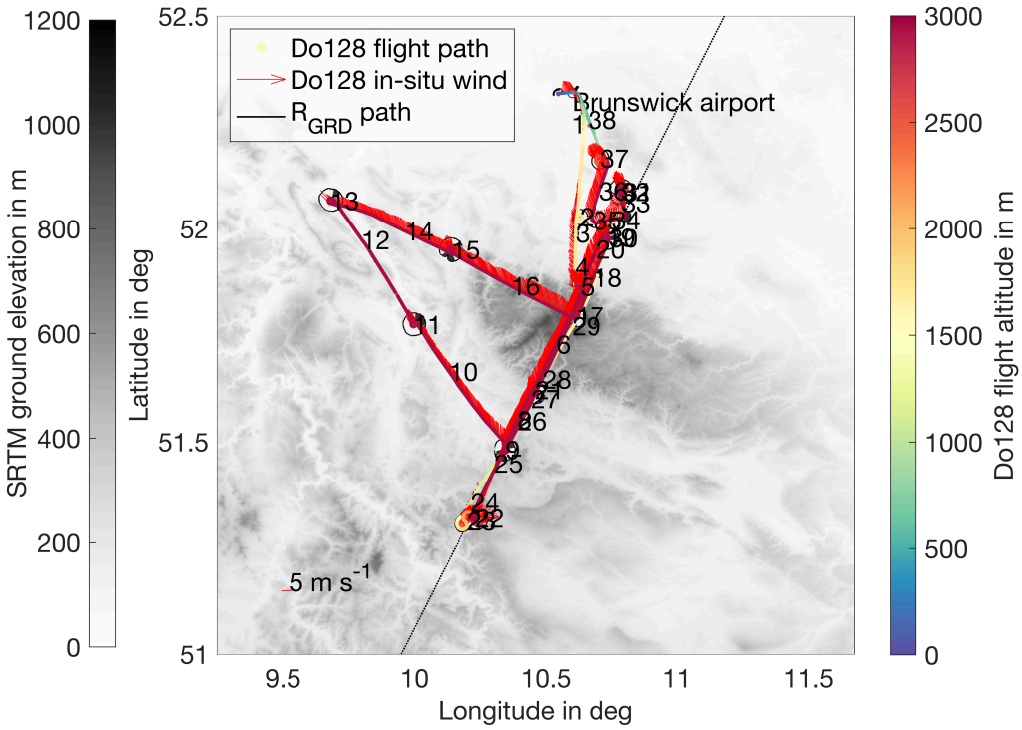

Fig. C.16.: Illustration of flight path for flight 2 (18 July 2019, flight time 06:15-08:58 UTC). Display characteristics as in Fig. C.15. Numbers correspond to the flight legs detailed in Tab. B.3, an overview of the ADL measured data is given in Fig. C.20.

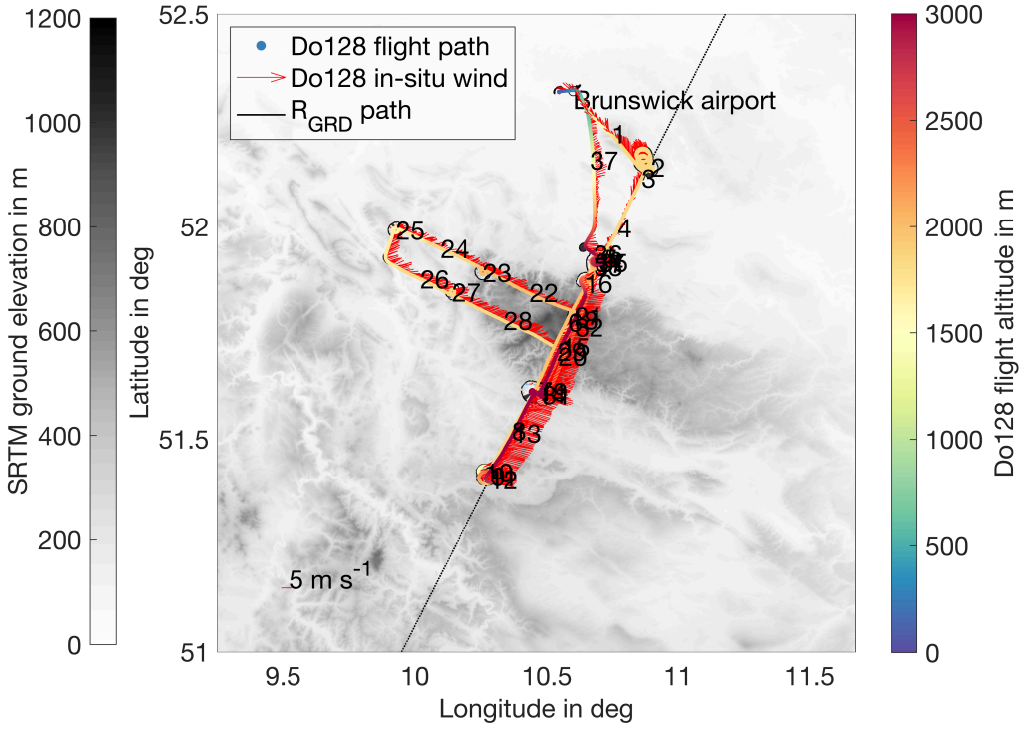

Fig. C.17.: Illustration of flight path for flight 3 (18 July 2019, flight time 12:30-15:14 UTC). Display characteristics as in Fig. C.15. Numbers correspond to the flight legs detailed in Tab. B.4, an overview of the ADL measured data is given in Fig. C.21.

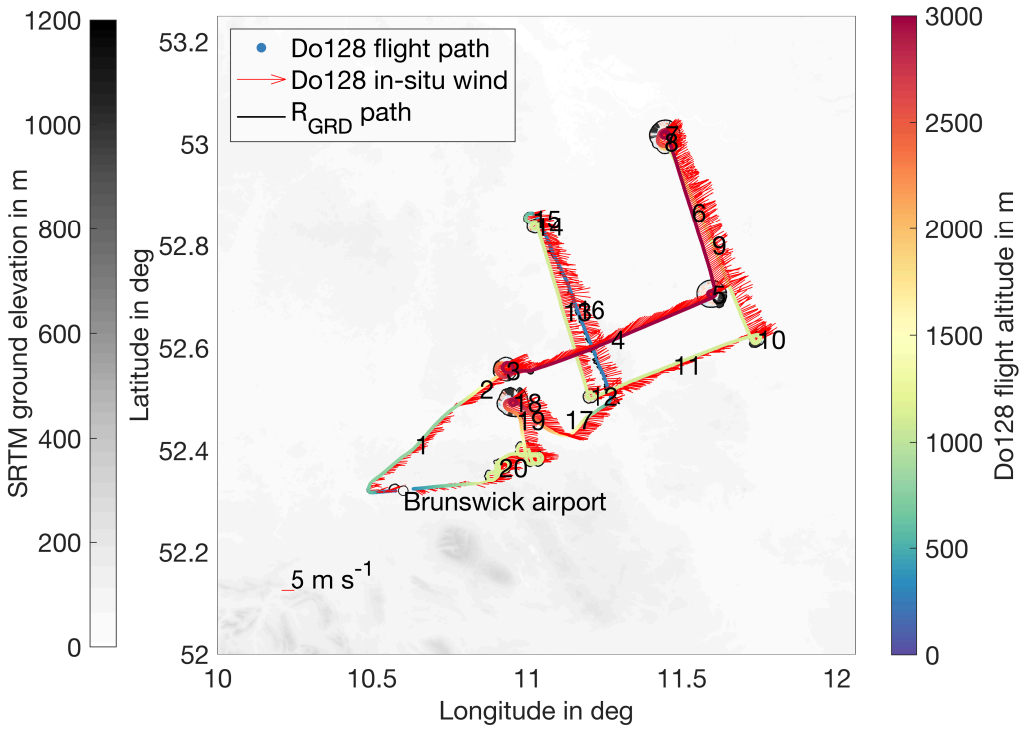

Fig. C.18.: Illustration of flight path for flight 4 (19 July 2019, flight time 09:40-11:38 UTC). Display characteristics as in Fig. C.15. Numbers correspond to the flight legs detailed in Tab. B.5, an overview of the ADL measured data is given in Fig. C.22.

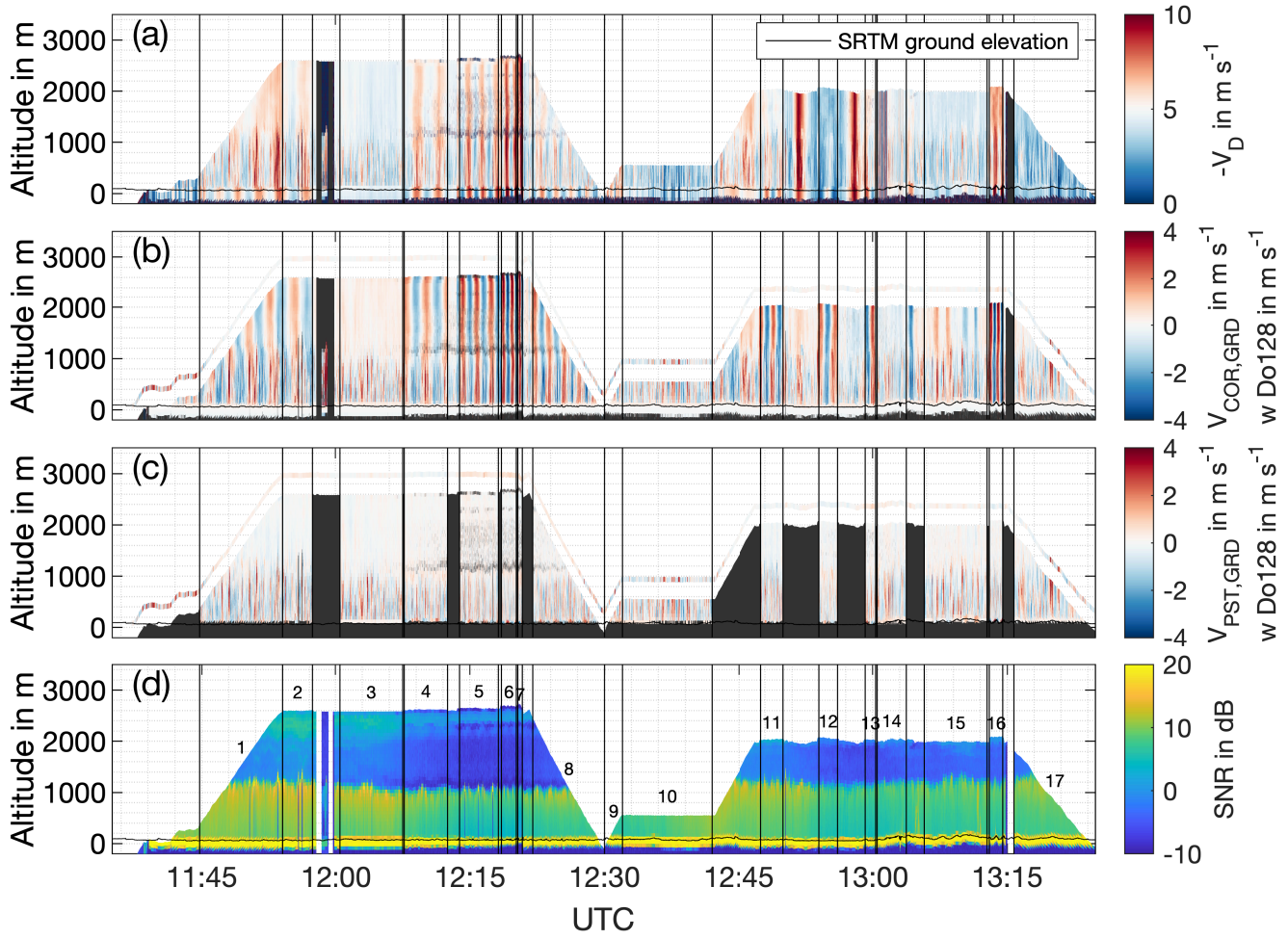

Fig. C.19.: Illustration of ADL measured data for flight 1 (10 July 2019, flight time 11:35-13:30 UTC). Displayed are: (a) Uncorrected lidar radial velocity measurement. (b) Refined motion corrected lidar radial velocity measurement (color coded curtain). Additionally, the aircraft vertical wind measurement is extended vertically by ±50 m and displayed at flight altitude as a color coded line. (c) Refined motion and horizontal wind corrected lidar measurement (color coded curtain). Aircraft vertical wind measurement as in panel above. (d) Lidar SNR. Numbers correspond to the flight legs detailed in Tab. B.2, vertical black lines show leg begin and end times. The ground elevation determined using a DEM is displayed as a black line in all panels. An overview of the horizontal flight path is given in Fig. C.15.

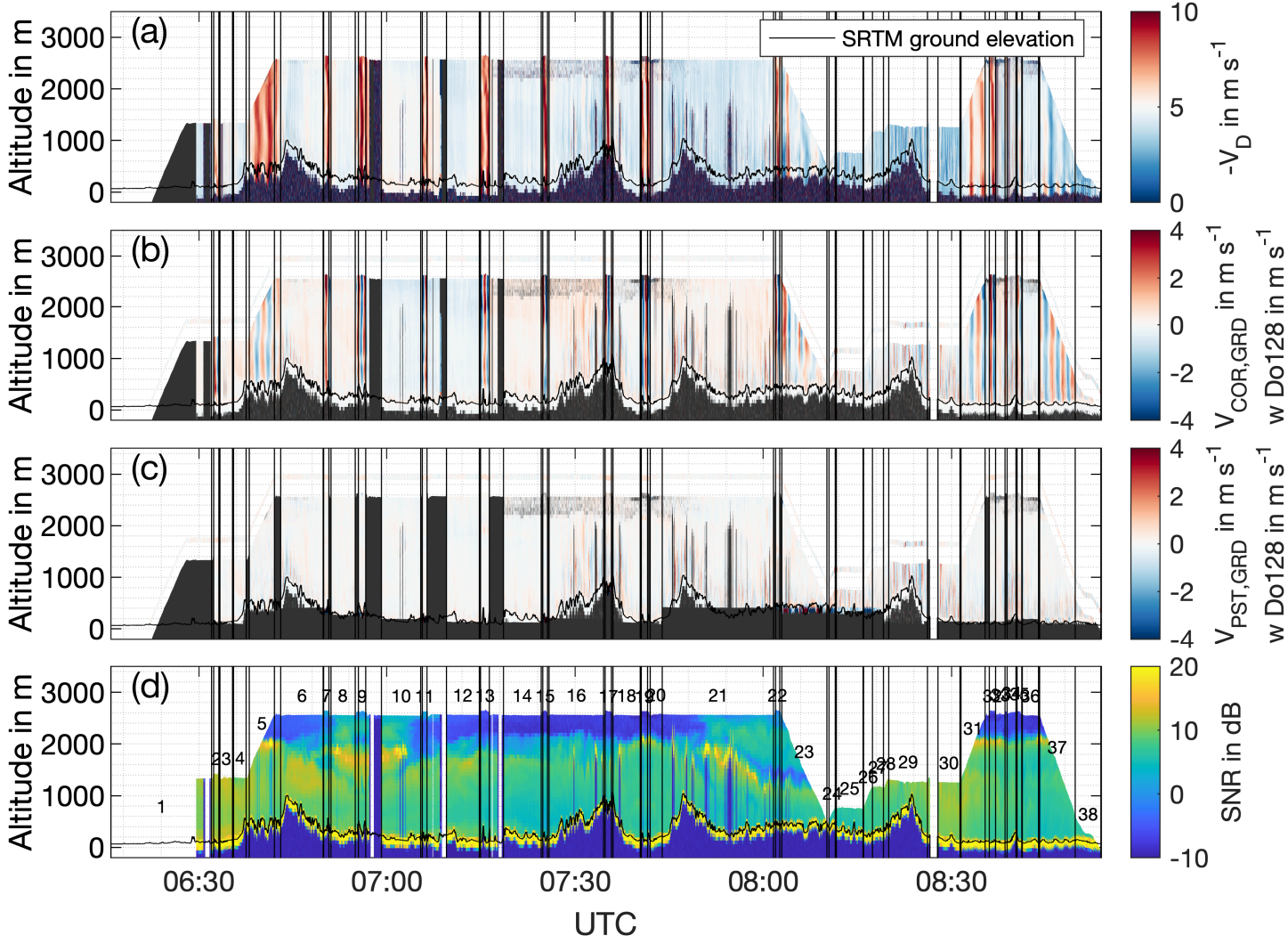

Fig. C.20.: Illustration of ADL measured data for flight 2 (18 July 2019, flight time 06:15-08:58 UTC). Display as in Fig. C.19. Numbers correspond to the flight legs detailed in Tab. B.3, vertical black lines show leg begin and end times. An overview of the horizontal flight path is given in Fig. C.16.

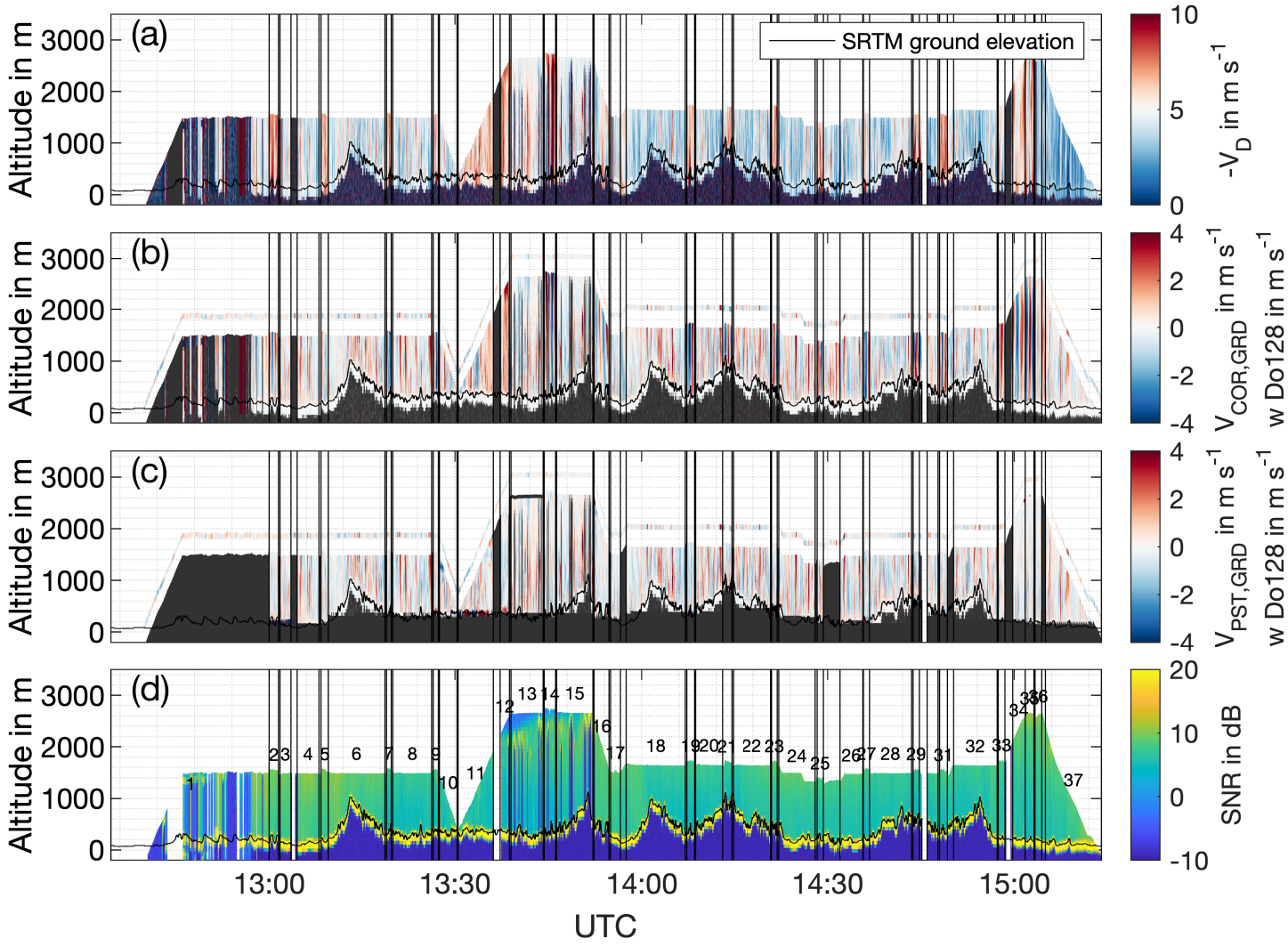

Fig. C.21.: Illustration of ADL measured data for flight 3 (18 July 2019, flight time 12:30-15:14 UTC). Display as in Fig. C.19. Numbers correspond to the flight legs detailed in Tab. B.4, vertical black lines show leg begin and end times. An overview of the horizontal flight path is given in Fig. C.17.

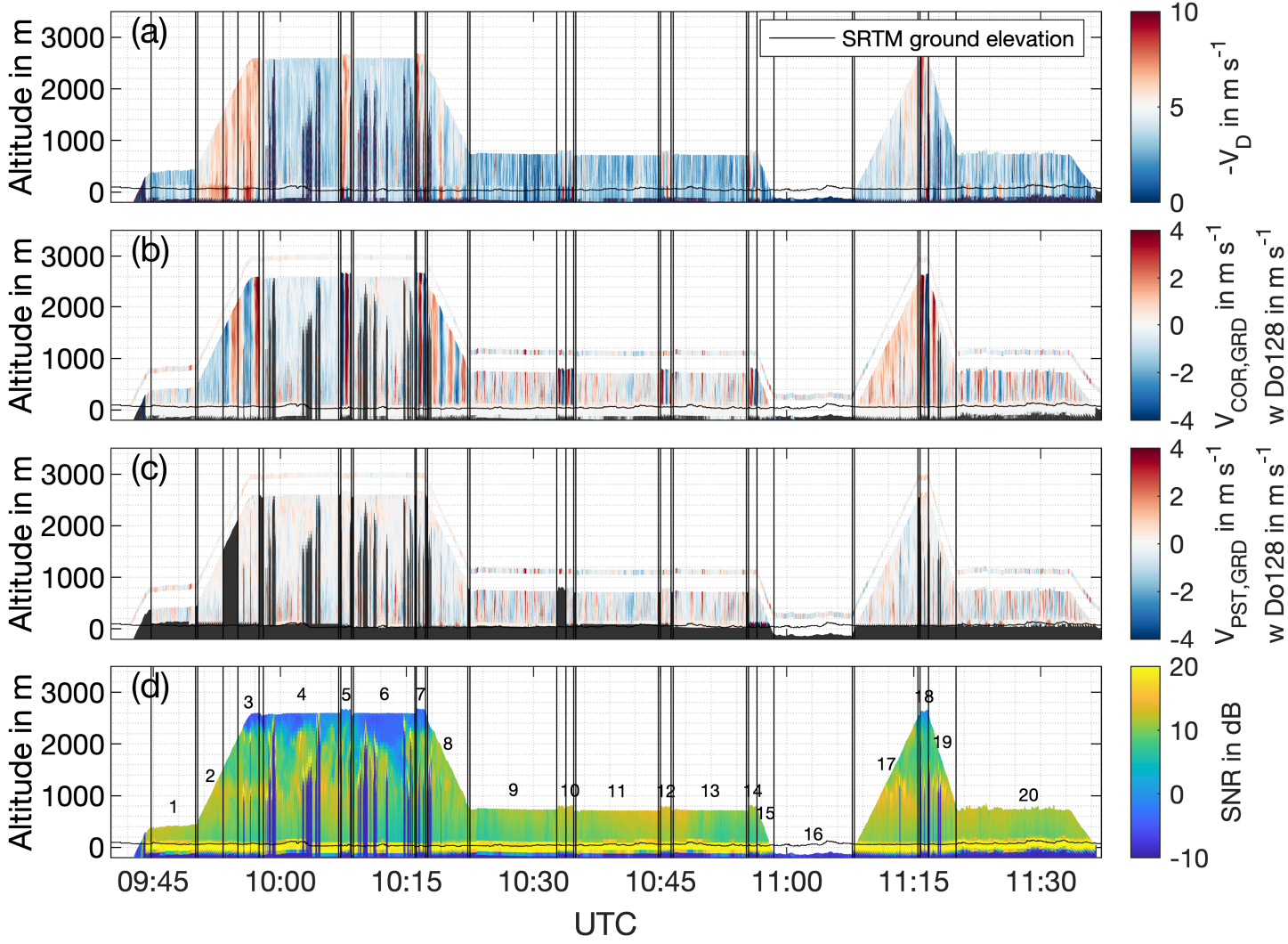

Fig. C.22.: Illustration of ADL measured data for flight 4 (19 July 2019, flight time 09:40-11:38 UTC). Display as in Fig. C.19. Numbers correspond to the flight legs detailed in Tab. B.5, vertical black lines show leg begin and end times. An overview of the horizontal flight path is given in Fig. C.18.

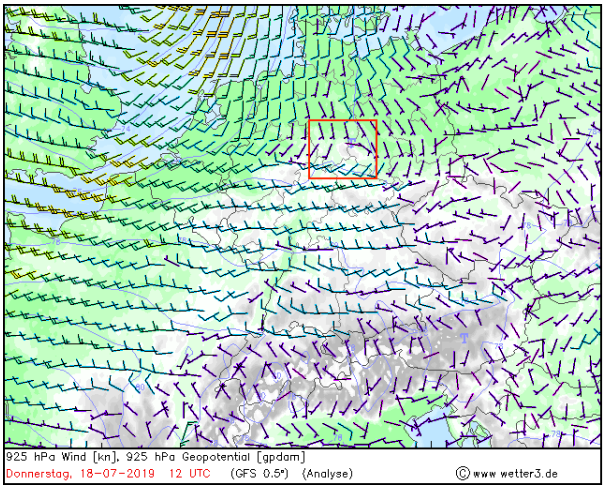

Fig. C.23.: Large-scale wind field at 925 hPa for 12 UTC, 18 June 2019 as modeled by the Global Forecast System. The location of the measurement area is marked by a red box. Figure obtained from www.wetter3.de.

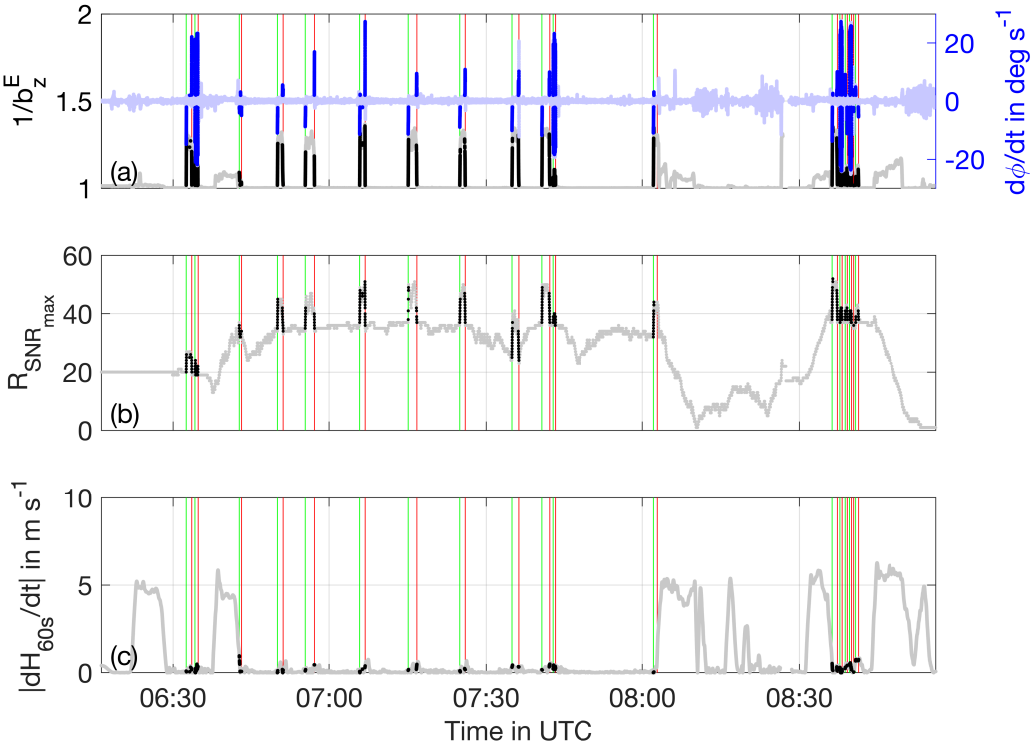

Fig. C.24.: Aircraft and lidar data used for time synchronization as a function of time for the morning flight on 18 July 2019. All data fulfilling the time synchronization criteria are marked as bold. (a) Aircraft roll data. (b) High frequency number of SNR_{max} range gate, nearest neighbor interpolated and quality controlled for cloud returns. (c) 60 s moving mean averaged vertical aircraft velocity. Beginning and end times of detected turn segments are shown as green and red vertical lines respectively. Note that PRY sequences show many short alternating roll oscillations per detected turn segment which are too close to be displayed.

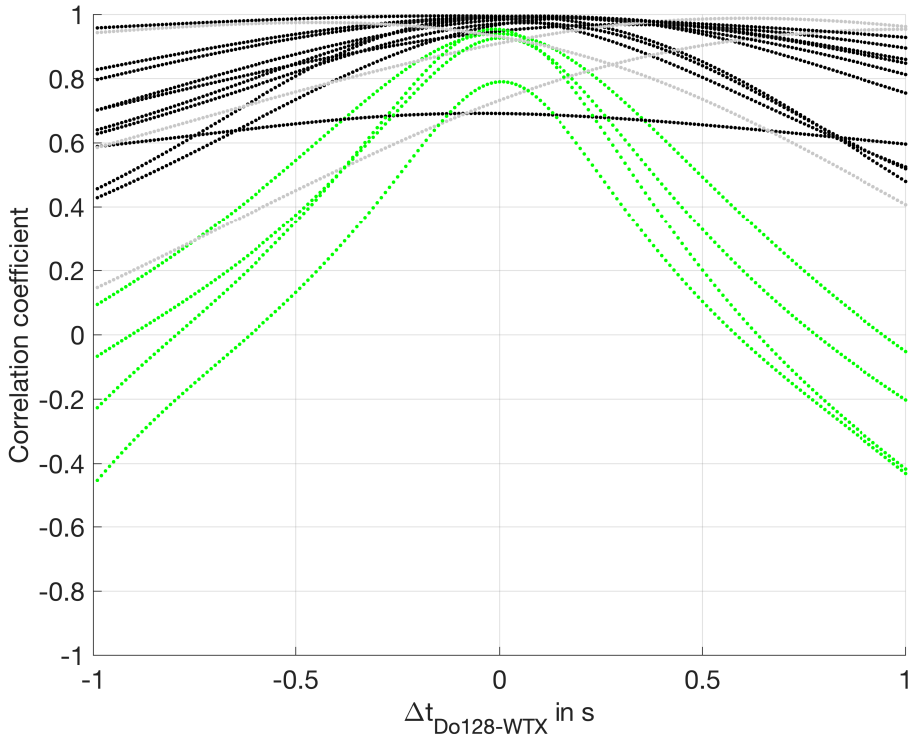

Fig. C.25.: Correlation coefficient of the variation of $1/b_z^E$ and the number of ground range gate as a function of time difference between internal Do128 time and internal WTX lidar time. Correlations are calculated for all time synchronization segments for the morning flight on 18 July 2019. The time difference $\Delta t_{Do128-WTX}$ is displayed for the synchronized data using a raw time difference $\Delta t_{Do128-WTX,raw} = 18.12\,\mathrm{s}$. Correlation coefficients obtained from PRY maneuvers are displayed in green, those obtained from other maneuvers (e.g. turns) are displayed in black. Segments which do not show a peak in the investigated interval are shown in gray.

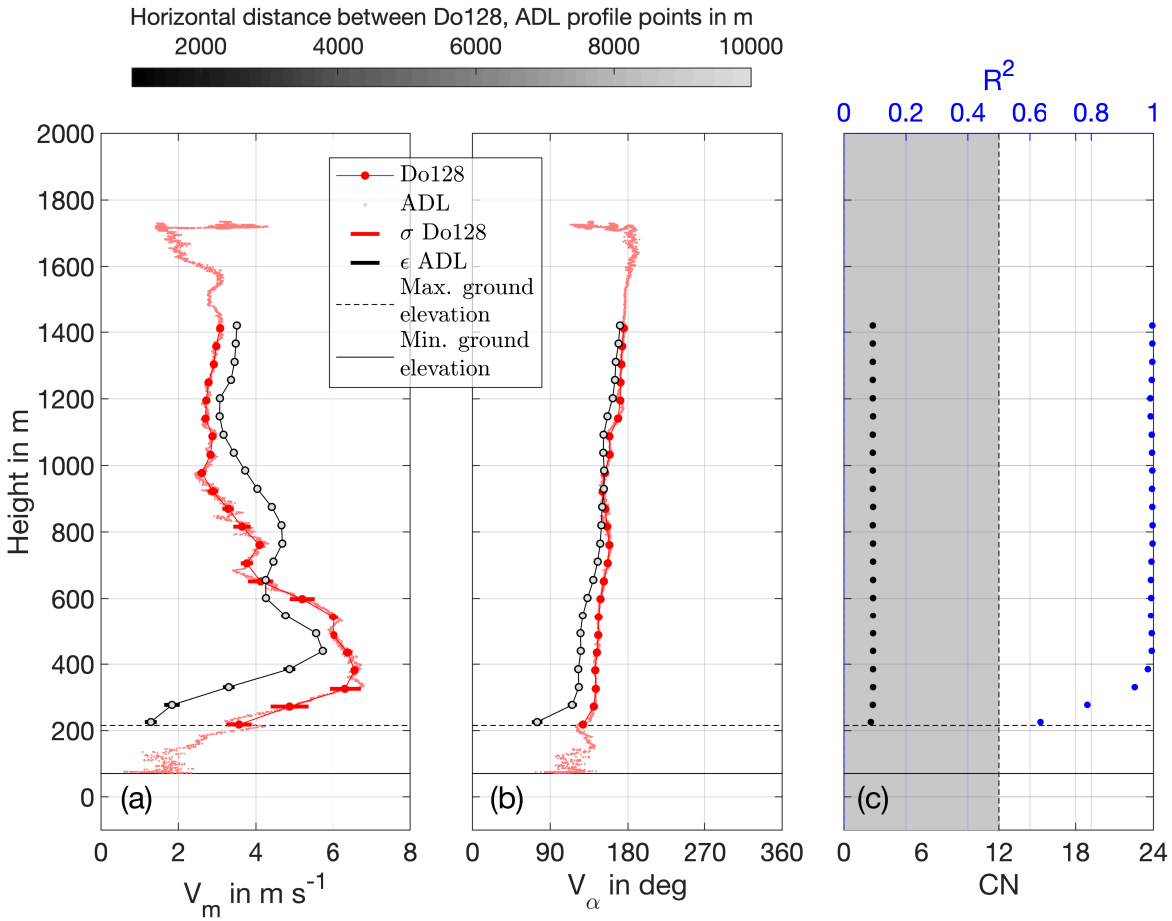

Fig. C.26.: Vertical profile of the horizontal wind retrieved with ADL from turn directly prior to PRY maneuver. (a) Wind speed. (b) Wind direction. (c) Quality control parameters R^2 and CN. The turn was conducted from 06:32:20 - 06:33:10 UTC. Additionally shown is the vertical profile of the horizontal wind obtained from Do128 in-situ measurements during a nearly straight climb-out from Brunswick airport towards the Harz region, 06:15:00 - 06:33:00 UTC. Please note that the distance covered by the aircraft during climb-out was approximately 60 km, therefore the lowest profile points differ by 60 km spatially and 15 min temporally, with decreasing distance towards the upper region of the profile. Due to the homogeneous wind field situation with little turbulence the ADL wind profile uncertainties are hidden behind the markers.

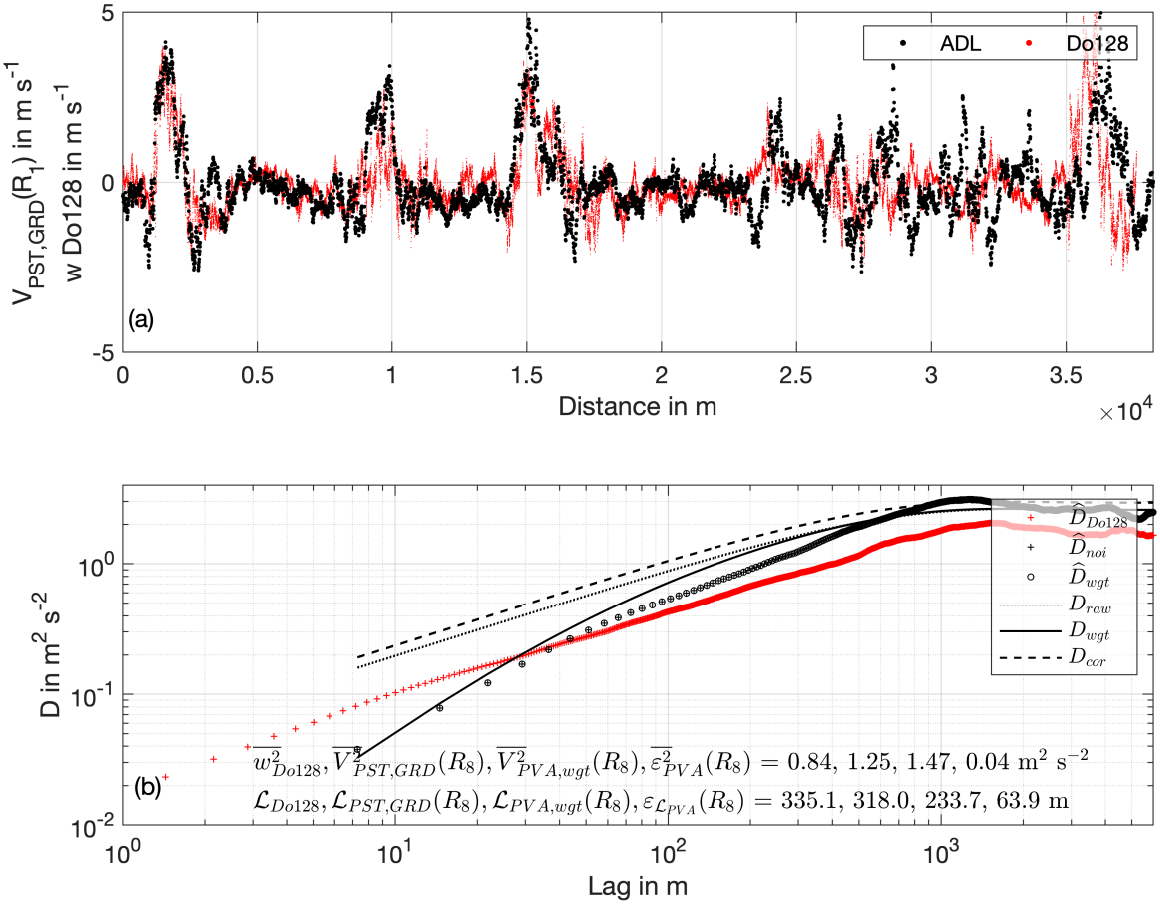

Fig. C.27.: Illustration of vertical wind measurements and structure functions used for PVA correction for top flight leg during flight 4 discussed in Sec. 6.7. (a) Displayed is the motion and wind corrected ADL measurement for the range gate R_1 conducted at 0.6 z_i (10:46:30-10:55:15 UTC). Further displayed is the in-situ Do128 vertical wind measurement, conducted simultaneously at 0.9 z_i, 400 m above the measurement of R_1. (b) Structure functions calculated using the vertical wind measurements and modeled using the modified azimuthal structure function technique (Frehlich et al., 2006). Labels according to the specifications given in the text, measured structure functions are marked as points and their symbols denoted with a hat, whereas idealized structure functions are marked as lines. The Do128 structure functions extends to smaller separations due to the higher measurement frequency of 100 Hz compared to 10 Hz for the ADL. Please note that the ADL measured variance and integral scale cannot be compared to in-situ measured values due to a vertical de-location of 400 m.

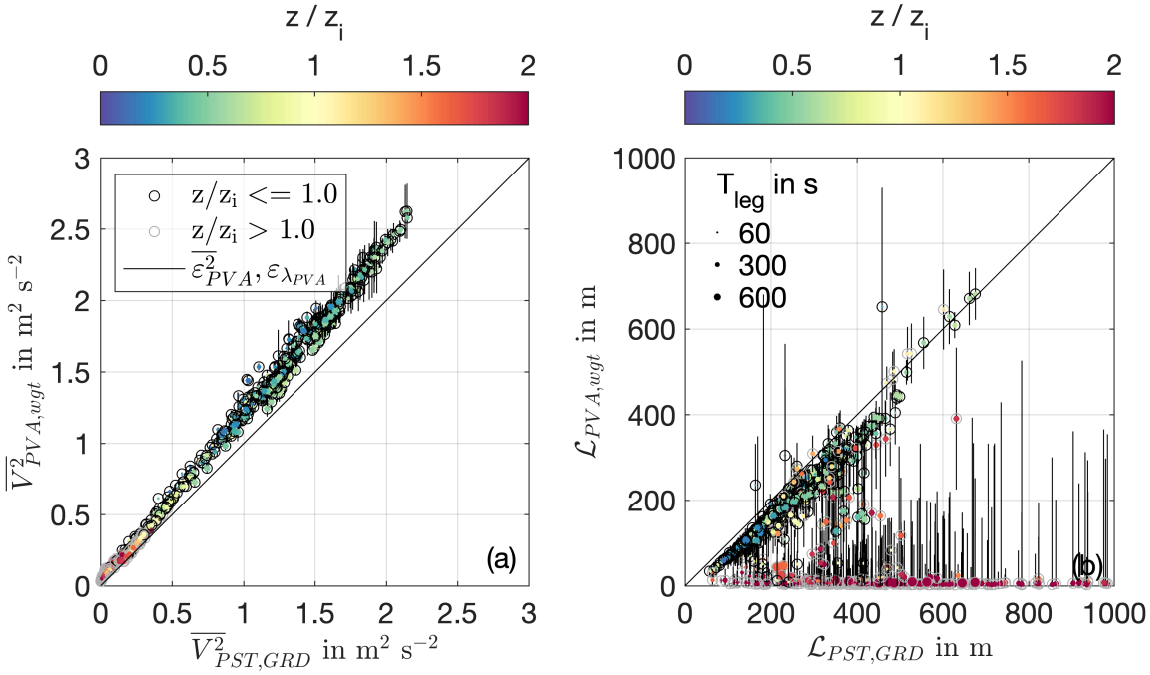

Fig. C.28.: Effect of PVA-correction procedure on (a) the variance and (b) the integral length scales. The PVA corrected values are shown in comparison to the values calculated directly using the definition. ADL measurements are displayed for all altitudes on all nadir legs with a minimum leg length of 60 s. The altitude of the individual measurements compared to the boundary layer height is color-coded, the von-Karman model is only expected to yield reliable results inside the boundary layer. The leg length is displayed through circle area. Deviations from the 1:1 line show the PVA-correction $\overline{\beta^2_{PVA}}$. For points circled in gray the PVA-correction is not applied in the final data due to their location above 1.0 z_i and the associated non-applicability of the von-Karman model. Vertical bars illustrate the uncertainty $\overline{\varepsilon^2_{PVA}}$ associated with the PVA correction procedure based on the raw von-Karman model applicability.

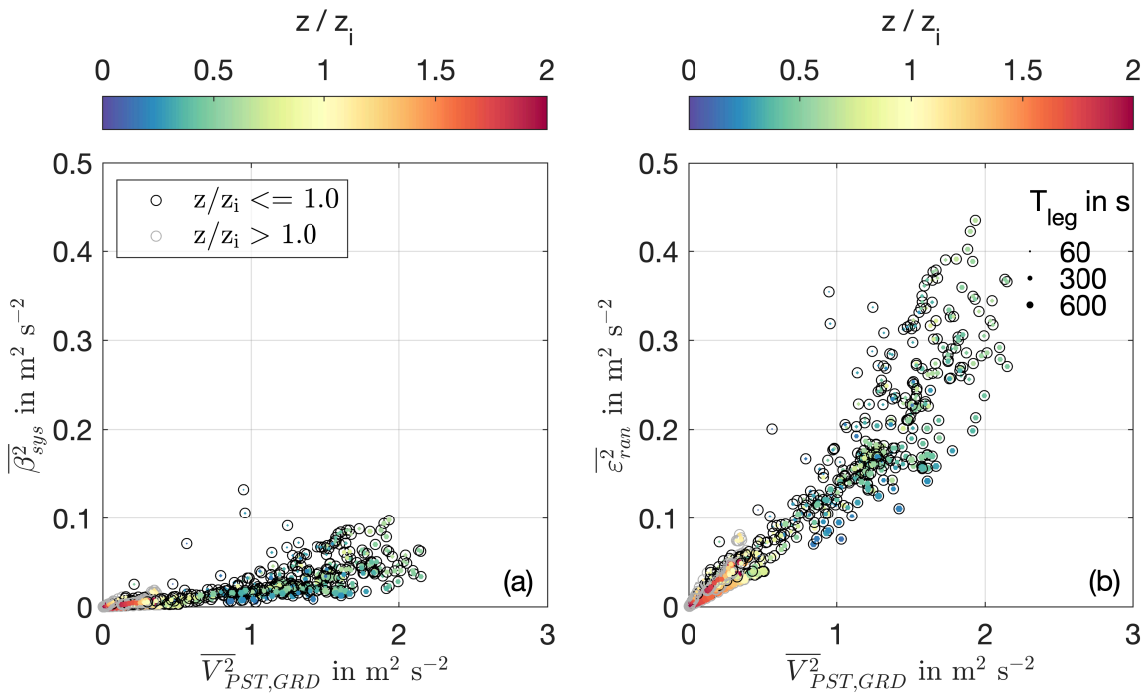

Fig. C.29.: Estimates of the systematic and random uncertainties due to the finite length of the investigated time series based on Lenschow et al. (1994). Shown are (a) the systematic offset (which is treated as an uncertainty of equal magnitude at the same time) and (b) the random uncertainty of the second order moments. ADL measurements are displayed for all altitudes on all nadir legs with a minimum leg length of 60 s. The leg length is displayed through circle area. The altitude of the individual measurements compared to the boundary layer height is color-coded.

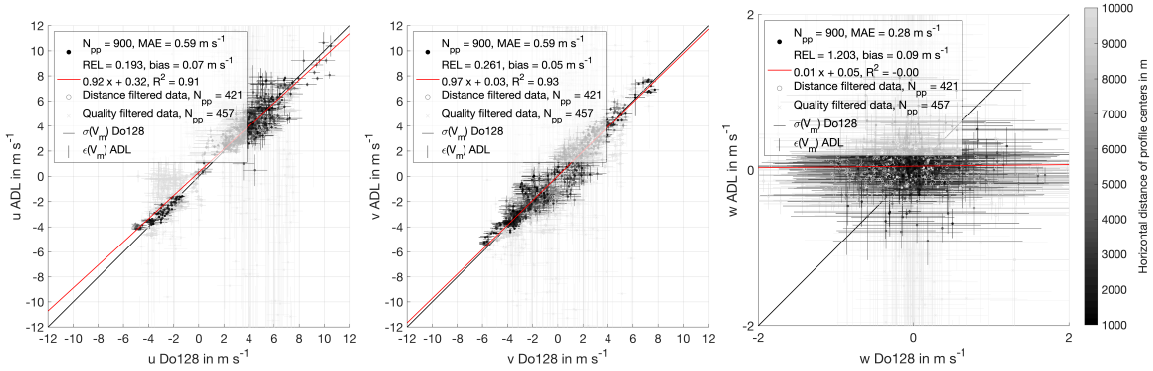

Fig. C.30.: Comparison of Do128 in-situ measured and ADL remotely sensed wind components u, v, w for all co-located profiles on all flights. Displayed quantities correspond to Fig. 6.21, except that the wind components are shown.

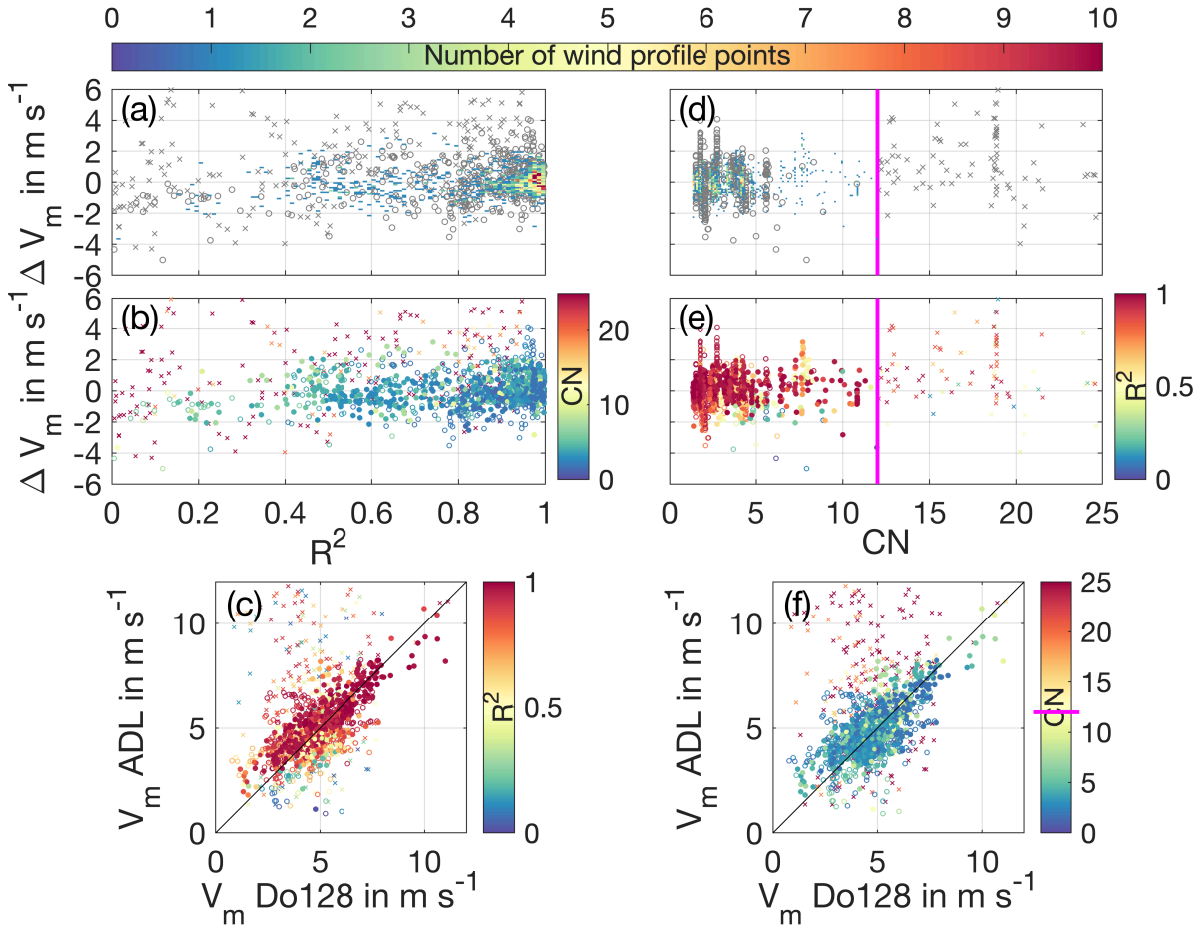

Fig. C.31.: Quality control criteria for all profiles on all flights, data as in fig 6.21. a) Color-coded histogram of number of occurrence for retrieval vs. in-situ difference and R^2. Crosses show values eliminated by quality control, hollow circles show values eliminated by distance control. b) Retrieval vs. in-situ difference and R^2, color-coded is CN. c) Do128 in-situ measured and ADL retrieved wind speed, color coded is R^2. Profile points which pass quality and distance control are displayed as color-coded full circles. d) Same as a), but for CN. e) Retrieval difference and CN, color coded is R^2. f) Same as c), but for CN. The magenta line indicates the CN quality control threshold applied.

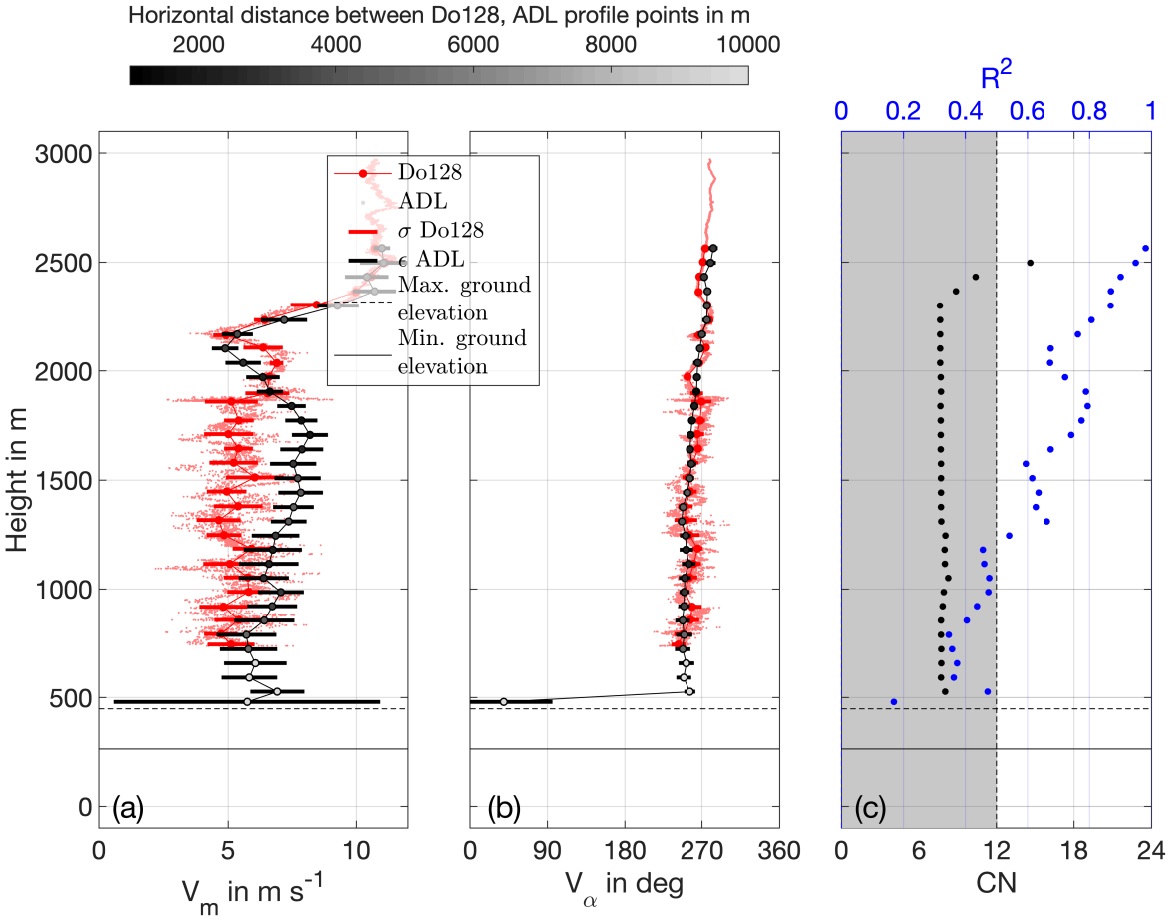

Fig. C.32.: Vertical profile of the horizontal wind retrieved with ADL from turn directly prior to PRY maneuver, with increased ADL uncertainty estimates due to inhomogeneous flow conditions. (a) Wind speed. (b) Wind direction. (c) Quality control parameters R^2 and CN. The turn was conducted from 13:37:15 - 13:38:45 UTC. Additionally shown is the vertical profile of the horizontal wind obtained from Do128 in-situ measurements during a close-by circling profile, 13:26:30 - 13:37:00 UTC.

Fig. C.33.: Cloud conditions encountered during flight 1 on 10 July 2019. Patches of less dense cumulus humilis (approx. 1/8) alternate with denser patches (approx. 4/8). At upper levels a thin cirrus layer is observed which thickened during the course of the flight. Picture taken at 11:56 UTC prior to the nadir transect investigated, location north of Wolfsburg, viewing direction north. The subsequent nadir transect analyzed in Sec. 6.5 was conducted towards the left side of the picture.

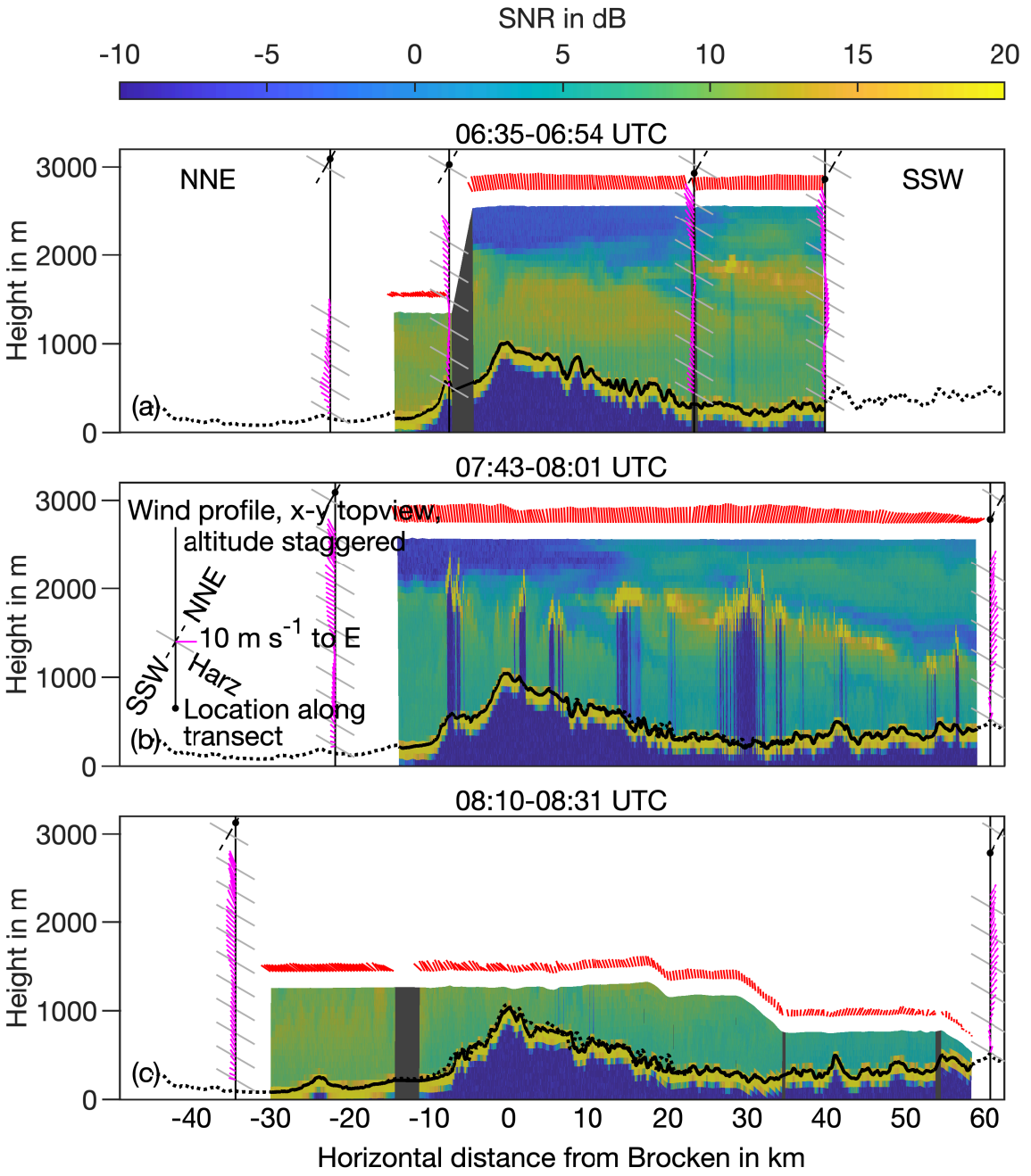

Fig. C.34.: Same as Fig. 6.27 but showing the SNR across the Harz mountain range during the morning flight on 18 July.

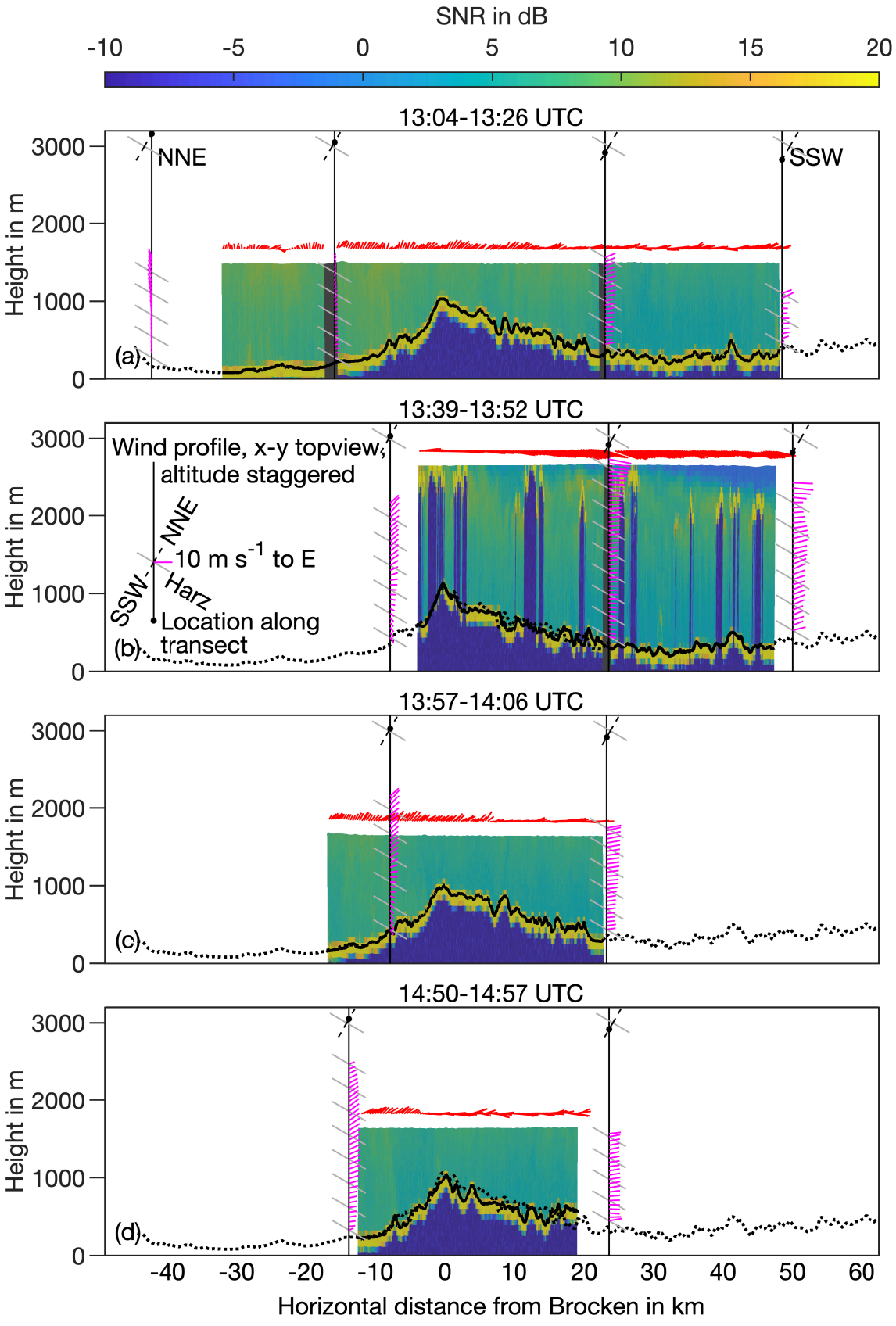

Fig. C.35.: Same as Fig. 6.28 but showing the SNR across the Harz mountain range during the afternoon flight on 18 July.

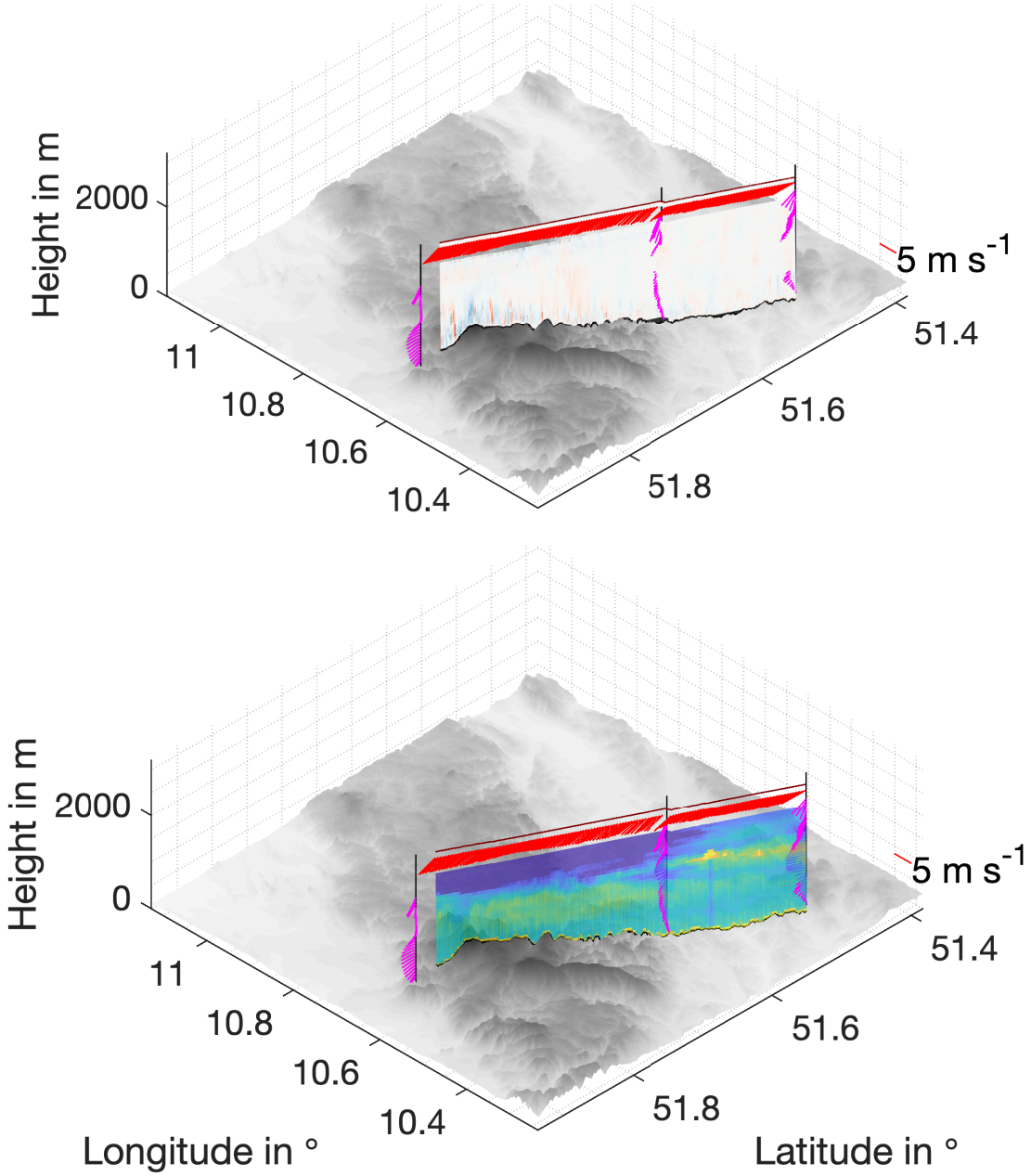

Fig. C.36.: 3-D illustration of the first transect across the Harz on the morning flight of 18 July 2019, 06:43-06:54 UTC. The viewing direction is towards south-east. Panel (a) shows the vertical wind, panel (b) shows the SNR. The associated color scales are as in Fig. 6.28 and Fig. C.35 respectively. Vertical lines illustrate the location of wind profiles, red vectors show wind direction and speed. Panel (b) is semi-transparent to show all wind profile vectors whereas panel (a) is not, thereby hiding some wind profile vectors.

Fig. C.37.: Wave-type cloud pattern observed in the free troposphere above the cumulus cloud field during flight 3 on 18 July 2019. Picture taken at 13:37 UTC on the southern side of the Harz, viewing direction southeast.

Danksagung

Ich beende diese Arbeit nach 4 Jahren mit einem lachenden und einem weinenden Auge, aber immerhin zwei Augen. Ein lachendes, weil jede weitere Seite dieser Arbeit definitiv eine zu viel gewesen wäre. Ein weinendes, weil viele Menschen diese Zeit zu einer schönen und besonderen gemacht haben, denen ich hier nur viel zu kurz Danke sagen will.
Mein besonderer Dank gilt Prof. Kottmeier für die Betreuung der Arbeit und Prof. Rapp für die Übernahme des Korreferats. Prof. Kottmeier hat mir seit Jahren immer wieder neue Möglichkeiten eröffnet um mich weiterzuentwickeln, sei es durch Auslandsaufenthalte oder Messkampagnen. Dank seinem Vertrauen konnte ich in dieser Arbeit meine eigenen Fragestellungen bearbeiten, wobei er stets unterstützend und interessiert zur Seite stand um jegliche Hindernisse aus dem Weg zu räumen. Prof. Rapp danke ich für das sofortige Interesse an der Arbeit und die Begutachtung, besonders im Angesicht der kurzfristigen Benachrichtigung und ungewissen Reisesituation.
Im Umfeld des IMK-TRO möchte ich Dr. Norbert Kalthoff für die Aufnahme in seine Arbeitsgruppe danken. Dank ihm konnte ich meine Ideen und Probleme stets offen diskutieren, außerdem danke ich ihm dafür jede einzelne Seite dieser Arbeit gelesen und korrigiert zu haben. Ebenso danke ich allen ehemaligen und aktuellen Mitgliedern dieser Arbeitsgruppe für die vielfältigen Diskussionen. Insbesondere möchte ich Dr. Andreas Wieser danken, von dem ich alles gelernt habe was ich über Messtechnik weiß, ohne ihn wären diese Arbeit und insbesondere die Messungen unmöglich gewesen. Gleiches gilt für Dr. Ulrich Corsmeier, der viele Stunden in die Koordination des Flugzeuglidarprojekts investiert und es entgegen aller Hindernisse stets gefördert hat. Das IMK-TRO und diese Arbeit würden nicht ohne die stets helfenden Hände der anderen Wissenschaftler, des Sekretariats und der Werkstatt funktionieren. Besonders Bernhard Deny und Samuel Haaß haben viel Mühe in dieses Projekt investiert. Meine inzwischen angesammelten Jahre am IMK-TRO bestanden glücklicherweise nicht nur aus Arbeit. Da ich äußerst gern Pausen mache gilt mein Dank auch hier allen Arbeitskollegen, die diese zu einer angenehmen Angelegenheit gemacht haben. Besonders Alberto, Jutta, Karmen, Lukas, Manuela, Simone und Shweta waren hierfür durch ihre offene und fröhliche Art über die Jahre unverzichtbar. Auch mein Studium wäre ohne Christian, Julia, Lisa und Hannah weniger schön verlaufen.

Zahlreiche Partner außerhalb des IMK-TRO haben ebenfalls zum Gelingen dieser Arbeit beigetragen. Für die Bereitstellung der alten und neuen LES-Simulationen danke ich der Arbeitsgruppe von Prof. Siegfried Raasch an der Universität Hannover und darin besonders Oliver Maas für die Simulation der neuen Windfelder. Ohne das IBUF-Team in Braunschweig, bestehend aus Dr. Thomas Feuerle, Rolf Hankers, Mark Bitter, Dr. Matthias Cremer und Helmut Schulz wäre kein einziger Datenpunkt aufgezeichnet worden. Ich danke für die stete Unterstützung des Lidar-Projektes auch in schwierigen Zeiten, sowie die Durchführung der Messflüge trotz teils Übelkeit erregender Flugzustände. Ich danke der Graduiertenschule GRACE für die unkomplizierte Finanzierung meines USA-Aufenthaltes. Besonders möchte ich mich bei Prof. Julie Lundquist für die Beherbergung in ihrer Arbeitsgruppe in Boulder bedanken, ebenso wie den zahlreichen Mitgliedern darin für die freundliche Aufnahme (besonders Nicola, Robert und Stephanie). Ebenso danke ich Dr. Alan Brewer sowie seiner Arbeitsgruppe für die interessanten Gespräche und hilfreichen Kommentare. Im Rahmen meines Besuchs in Wyoming danke ich herzlich Dr. Samuel Haimov und seiner Frau für die tolle Gastfreundschaft. Als letztes Danke ich dem LSV Albgau und den engagierten Leuten darin, die mir einen Traum erfüllt und mich das Fliegen gelehrt haben.

Auch für die Gestaltung meiner restlichen Lebenszeit während der letzten Jahre möchte ich mich herzlich bei vielen Leuten bedanken. Auch wenn die Schulzeit schon ein Weilchen vorbei ist danke ich meinen Klassenkameraden für die bis heute erhaltenen Freundschaften, besonders Carolin, Daniel, David, Gloria, Katrin, Markus, Petra und Sebastian. Mein herzlicher Dank für den Spaß in und außerhalb des Studiums gilt auch meinen Studienkollegen Anton, Isabel, Lukas, Malte, Max, Michael, Moritz, Rainer, Robert und Simon. Auch viele andere Menschen machen Karlsruhe zu einer schöneren und lebenswerteren Stadt durch Ausflüge und Volleyball, darunter Gabi, Felix und Stefan sowie Alex, Andi, David, Simon und natürlich Raimund als Kapitän. Für die spaßige Einführung in die Kultur, Küche und Philosophie der Deutsch-Französischen Freundschaft danke ich allen aus dieser Kommune, ganz besonders Annabel, Ariane, Fred, Leo, Josquin sowie Jude und Nicolas. Einige wenige haben es geschafft mich in einer Doppelfunktion als WG-Mitbewohner und teils sogar Arbeitskollege auszuhalten. Für diese stets mit Bravour gemeisterte Aufgabe danke ich Chloé, Maren und Andreas. Eventuell ist es ganz gut, dass ich bisher nur tages- oder wochenweise mit Jan zusammen gewohnt habe, denn so gebührt ihm mein ganz besonderer Dank als mein unverzichtbarer, bester Freund in allen Lagen. Jan wiederum wäre nicht der gleiche ohne die wunderbare Begleitung an seiner Seite, Denise, der unabhängig von ihm ebenso mein Dank gebührt.

Meine Familie ist dafür verantwortlich, dass ich immer meinen Interessen nachgehen und mich dabei auf Rückhalt verlassen konnte. Nach Sachsen, besonders an Familie Thomas und ihre

Erweiterungen, geht mein Dank für die vielen schönen Besuche und Feste. Ebenso möchte ich mich bei meinen Großeltern für die immer gewährte Unterstützung bedanken. Mein größter Dank geht natürlich an meine Familie und darin besonders an meine Mutter, die den Laden zusammen und in Schuss hält. Durch meine Geschwister Richard und Luise wird Familienzeit immer zu etwas wunderbarem, es ist mir eine große Freude euer großer Bruder zu sein.
Vielen Dank Franziska, du hast jeden Tag dieser Arbeit zu einem besseren und besonderen gemacht.

Acronyms

ADL	Airborne Doppler Lidar
ADLS	Airborne Doppler Lidar Simulator
ADR	Airborne Doppler Radar
AGL	Above Ground Level
ARD	Along-Right wing-Down reference system
AVAD	Airborne Velocity Azimuth Display technique
AVG	Along-track AVeraGing distance
BL	(Atmospheric) Boundary Layer
BLA	Sector BLAnking
DAWN1	ADL on NASA DC-8 aircraft
DAWN2	ADL on UC-12B aircraft
DEM	Digital Elevation Model
DIAL	DIfferential Absorption Lidar
DLR	Deutsches Zentrum für Luft- und Raumfahrt
DLR 2μm	ADL 2μm onboard DLR Falcon 20 aircraft
DLR A2D	Direct detection ADL onboard DLR Falcon 20 aircraft
DNS	Direct Numerical Simulation
Do128	Dornier 128-6 research aircraft of the TU Brunswick, call-sign D-IBUF
ECEF	Earth-Centered, Earth-Fixed reference system
ELE	Scan ELEvation angle
ENU	East-North-Up reference system
ESRL	Earth System Research Laboratory
FFT	Fast Fourier Transform
FME	Lidar Frequency of MEasurement

FWHM	Full Width at Half Maximum
GEO	GEOdetic reference system
GrOAWL	Green Optical Autocovariance Wind Lidar onboard NASA WB-57 aircraft
GS	Grid Scale
HGT	HeiGhT used for retrieval
IAS	Indicated Air Speed
IFR	Instrument Flight Rules
INS	Inertial Navigation System
LASER	Light Amplification by Stimulated Emission of Radiation
LES	Large Eddy Simulation
LEV	Vertical averaging LEVel
LIDAR	LIght Detection And Ranging
LIVE	ADL onboard ATR-42 aircraft
LOS	Line-Of-Sight
LSQ	Least-SQuares
M2D	ADL onboard Twin-Otter aircraft
MSL	Mean Sea Level
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
NVS	Navier-Stokes
NWP	Numerical Weather Prediction
OAWL	Optical Autocovariance Wind Lidar onboard NASA WB-57 aircraft
ONERA	Office National d'Études et de Recherches Aérospatiales
P3DWL	ADL onboard P3-D aircraft
PALM	PArallelized Large-eddy simulation Model
PRF	Pulse Repetition Frequency

PRY	Pitch, Roll, Yaw
PVA	Pulse Volume Averaging
RANS	Reynolds Averaged Navier Stokes
RET	Standard wind profile RETrieval strategy
ROT	Scan ROTation speed
SAFIRE	Service des Avions Francais Instrumentés pour la Recherche en Environnement
SGS	Sub-Grid Scale
SRTM	Shuttle Radar Topography Mission
STP	Standard system SeTuP
TODWL	ADL onboard Twin-Otter aircraft
UTC	Coordinated Universal Time
VAD	Velocity Azimuth Display technique
VFR	Visual Flight Rules
VVP	Volume Velocity Processing
WGS84	World Geodetic System WGS84 reference ellipsoid
WND	Background WiND speed case
WTX	Lockheed Martin Coherent Technologies 1.6 μm WindTracer Doppler lidar

List of Symbols

Symbol	Description	Unit
	General sub- and superscript conventions	
$[]^{A}$	Superscript for quantities in aircraft reference system ARD	
$[]^{C}$	Superscript for quantities in Earth-centered, Earth-fixed reference system ECEF	
$[]^{E}$	Superscript for quantities in ground reference system ENU	
$[]^{G}$	Superscript for quantities in geodetic reference system GEO	
$[]^{LES}$	Superscript for individual input-truth quantities in ADLS	
$[]^{R}$	Superscript for retrieved quantities in ADLS	
$[]^{T}$	Superscript for averaged input-truth quantities in ADLS	
$[]_{GRD}$	Subscript for quantities after using ground-return for refined motion correction	
$[]_{PST}$	Subscript for quantities after post-processing for horizontal wind contribution removal	
$[]_{RAW}$	Subscript for quantities without application of ground-return for refined motion correction	
$[]'$	Transpose for matrices	
q'	Turbulent deviation of q from mean for scalar quantities	
$\overline{[]^{1}}$	Notation for first-order moment	
$\overline{[]^{2}}$	Notation for second-order moment	
	Transformation and rotation matrices	
$\mathbb{H}$	Heading rotation matrix	
$\mathbb{P}$	Pitch rotation matrix	
$\mathbb{R}$	Roll rotation matrix	

Symbol	Description	Unit
$\mathbb{T}^{AE}$	Transformation matrix from aircraft reference frame ARD to East-North-Up reference frame ENU	
$\mathbb{T}^{CE}$	Transformation matrix from Earth-centered, Earth-fixed reference frame ECEF to East-North-Up reference frame ENU	
	General variables	
A	Area	m^2
H	Height	m
N	Number of measurements	
Ω_g	Earth's angular velocity	$\mathrm{rad\,s}^{-1}$
$\mathcal{R}$	(Auto-)covariance	Depending on quantity
TKE	Turbulent kinetic energy	$\mathrm{m}^2\,\mathrm{s}^{-2}$
Θ	Potential temperature	K
t	Total measurement time	s
$\boldsymbol{b}$	Beam direction unit vector	
c	Speed of light	$\mathrm{m\,s}^{-1}$
f	Frequency	Hz
g	Gravitational acceleration	$\mathrm{m\,s}^{-2}$
$\mathcal{L}$	Integral length scale	m
$\mathcal{T}$	Integral time scale	s
λ_m	Wavelength of the spectral maximum	m or s
λ	Wavelength	m or s
ν	Kinematic viscosity coefficient	$\mathrm{m}^2\,\mathrm{s}^{-1}$
ω	Angular frequency	$\mathrm{rad\,s}^{-1}$
p	Pressure	hPa
ρ_a	Density of dry air	$\mathrm{kg\,m}^{-3}$
ρ	(Auto-)correlation	
τ	Lag	m or s
t	Time	s
z_i	Boundary layer height	m
δ_{ij}	Kronecker delta, $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ for $i \neq j$	

Symbol	Description	Unit
ε_{ijk}	Levi-Civita symbol, $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$; $\varepsilon_{132} = \varepsilon_{213} = \varepsilon_{321} = -1$; $\varepsilon_{ijk} = 0$ otherwise	
$\overline{\sigma^1}$	Measured standard deviation of first order moment	Depending on quantity
$\overline{\mu^1}$	First-order moment	Depending on quantity
$\overline{\mu^2}$	Second-order moment	Depending on quantity
	Lidar parameters	
ΔR	Effective range resolution	m
Δh	Movement of range gate during measurement	m
Δp	Range gate length	m
Δr	Lidar pulse length	m
E	Energy level	J
I_n	Gaussian pulse weighting function	
Φ	Scan angle	rad
P	Signal power	W
R_{GRD}	Number of ground range gate	
$R_{SNR_{max}}$	Number of range gate with maximum SNR	
R_{dis}	Distance to the center of the range gate	m
R_n	Number of n-th range gate	
SNR	Signal-to-noise ratio	
SpG	Samples per range gate	
α	Extinction coefficient	$\mathrm{sr}^{-1}\,\mathrm{m}^{-1}$
β	Backscatter coefficient	$\mathrm{sr}^{-1}\,\mathrm{m}^{-1}$
η	Detector efficiency parameter	
f_{spl}	Lidar sampling rate of signal detector	Hz
$\hbar$	Planck constant	J s
λ_L	Lidar wavelength	m
n_R	Range gate number index	
r	Distance along beam	m
σ_p	1/e width of lidar pulse	s
v_{pulse}	Pulse weighted radial velocity	$\mathrm{m\,s}^{-1}$

Symbol	Description	Unit
	Aircraft parameters	
TAS	True air speed	$\mathrm{m\,s^{-1}}$
X	Cartesian x-coordinate	m
Y	Cartesian y-coordinate	m
Z	Cartesian z-coordinate	m
a	Equatorial radius of WGS84 ellipsoid	m
b	Polar radius of WGS84 ellipsoid	m
h_g	Geodetic height	m
λ_g	Geodetic longitude	deg
$\boldsymbol{\omega}$	Aircraft rotation rate vector	$\mathrm{rad\,s^{-1}}$
$d\phi/dt$	Aircraft roll rate	$\mathrm{rad\,s^{-1}}$
ϕ	Aircraft roll angle	rad
ψ	Aircraft heading angle	rad
$\boldsymbol{r}$	Lidar moment arm, distance IMU to lidar	m
θ	Aircraft pitch angle	rad
φ_g	Geodetic latitude	deg
	Velocity parameters	
$S(f)$	Variance density in spectral decomposition	$\mathrm{m^2\,s^{-1}}$
V_α	Wind direction in retrieval volume, from ADL retrieval, ADLS or aircraft measurement	deg
V_m	Vector averaged wind speed in retrieval volume, from ADL retrieval, ADLS or aircraft measurement	$\mathrm{m\,s^{-1}}$
$\tilde{V}_m$	Scalar averaged wind speed in retrieval volume from ADLS	$\mathrm{m\,s^{-1}}$
u	Horizontal wind component towards east	$\mathrm{m\,s^{-1}}$
v_{COR}	Motion corrected radial Doppler velocity	$\mathrm{m\,s^{-1}}$
v_D	Measured radial Doppler velocity	$\mathrm{m\,s^{-1}}$
$\boldsymbol{v}_L$	Velocity of lidar	$\mathrm{m\,s^{-1}}$
$\boldsymbol{v}_a$	Velocity of aircraft	$\mathrm{m\,s^{-1}}$
$\boldsymbol{v}_p$	Velocity of scattering particles	$\mathrm{m\,s^{-1}}$
$\boldsymbol{v}$	Velocity vector	$\mathrm{m\,s^{-1}}$
$\overline{v_r^2}$	Raw radial velocity variance, second-order moment	$\mathrm{m^2\,s^{-2}}$
v	Horizontal wind component towards north	$\mathrm{m\,s^{-1}}$

Symbol	Description	Unit
$\overline{w^2}$	Variance of the vertical wind component, second-order moment	$\mathrm{m^2\,s^{-2}}$
w	Vertical wind component	$\mathrm{m\,s^{-1}}$
	Wind profile retrieval paramaters	
CN	Condition number	
$\mathbb{C}_m$	Covariance matrix of model parameters	
$\mathbb{G}^{-g}$	General inverse of beam pointing matrix	
$\mathbb{G}$	Beam pointing matrix	
$\mathbb{I}$	Identity matrix	
N_{corr}	Sample size corrected for along beam correlation	
N_{eff}	Effective sample size corrected for correlation among residuals	
R^2	Coefficient of determination	
X_a	Along track averaging distance	m
X_c	Across track averaging distance	m
X_d	Altitude dependent scan diameter	m
λ_S	Singular values of beam pointing matrix	
$\overline{\mu_d^2}$	Data variance, estimated by residual variance	$\mathrm{m^2\,s^{-2}}$
m	Number of model parameters, retrieved wind speed components	
ρ_{res}	Correlation among residuals from neighboring range gates	
	Quality metrics	
MAE	Mean absolute error	$\mathrm{m\,s^{-1}}$
N_n	Normalized number of wind profile points (by standard system setup and retrieval strategy)	
N_{pp}	Number of wind profile points	
N_p	Number of wind profiles	
REL	Relative root mean square error, normalized with the wind speed	
$RMSE$	Root mean square error	$\mathrm{m\,s^{-1}}$
	Nadir retrieval quantities	
D	Structure function	$\mathrm{m^2\,s^{-2}}$

Symbol	Description	Unit
Γ	Gamma function	
K	Bessel function	
L_0	Outer scale of turbulence	m
Λ_D	Transversal component of von-Karman model	
Λ	Longitudinal component of von-Karman model	
s	Separation distance	m
	LES quantities	
E_{GS}	Grid scale turbulent kinetic energy	$\mathrm{m^2\,s^{-2}}$
E_{SGS}	Sub-grid scale turbulent kinetic energy	$\mathrm{m^2\,s^{-2}}$
Q_Θ	Source term for heat	$\mathrm{K\,s^{-1}}$
$\boldsymbol{U_g}$	Geostrophic wind vector	$\mathrm{m\,s^{-1}}$
ϵ	Dissipation rate	$\mathrm{m^2\,s^{-3}}$
	Uncertainties	
$\Delta \mathrm{V}^{w'}_{phy}$	Physically parameterized supra-scan-volume uncertainty estimation	$\mathrm{m\,s^{-1}}$
$\Delta \mathrm{V}^{w'}_{proj}$	LES-determined supra-scan-volume flow inhomogeneity contribution to wind profiling error	$\mathrm{m\,s^{-1}}$
$\mathrm{UNC_{cov}}$	Average value of raw covariance matrix based sub-scan-volume uncertainty estimation	$\mathrm{m\,s^{-1}}$
$\mathrm{UNC_{dir}}$	Average value of LES-determined supra-scan-volume effective sample size corrected covariance matrix based sub-scan-volume uncertainty estimation	$\mathrm{m\,s^{-1}}$
$\mathrm{UNC_{eff}}$	Average value of effective sample size corrected covariance matrix based sub-scan-volume uncertainty estimation	$\mathrm{m\,s^{-1}}$
$\mathrm{UNC_{phy}}$	Average value of physically parameterized supra-scan-volume uncertainty estimation and effective sample size corrected covariance matrix based sub-scan-volume uncertainty estimation	$\mathrm{m\,s^{-1}}$
$\mathrm{UNC_{sup}}$	Average value of physically parameterized supra-scan-volume uncertainty estimation	$\mathrm{m\,s^{-1}}$
ε_{V_α}	Uncertainty associated with retrieved wind direction	deg

Symbol	Description	Unit
ε_{V_m}	Covariance matrix based sub-scan-volume uncertainty estimation	$\mathrm{m\,s^{-1}}$
ε_u	Uncertainty associated with retrieved u wind component	$\mathrm{m\,s^{-1}}$
ε_v	Uncertainty associated with retrieved v wind component	$\mathrm{m\,s^{-1}}$
ε_w	Uncertainty associated with retrieved w wind component	$\mathrm{m\,s^{-1}}$
$\boldsymbol{\varepsilon_a}$	Uncertainty in platform ground velocity, including due to rotation	$\mathrm{m\,s^{-1}}$
$\overline{\varepsilon^1}$	Random uncertainty associated with first-order moment	Depending on quantity
$\overline{\beta^1}$	Systematic offset uncertainty associated with first-order moment	Depending on quantity
$\overline{\varepsilon^2}$	Random uncertainty associated with second-order moment	Depending on quantity
$\overline{\beta^2}$	Systematic offset uncertainty associated with second-order moment, can be estimated through existing theory for some	Depending on quantity
$\overline{\varepsilon^1_{HC}}$	Random uncertainty associated with first-order moment due to contribution of horizontal wind to measured vertical wind, e.g. due to aircraft pitch and roll	$\mathrm{m\,s^{-1}}$
$\overline{\varepsilon^1_{ac}}$	Random uncertainty associated with first-order moment due to aircraft motion correction	$\mathrm{m\,s^{-1}}$
$\overline{\varepsilon^1_b}$	Random uncertainty associated with first-order moment due to beam pointing-angle calibration	$\mathrm{m\,s^{-1}}$
$\overline{\varepsilon^1_{ran}}$	Random uncertainty associated with first-order moment due to substitution of ensemble mean with time series mean	$\mathrm{m^2\,s^{-2}}$
$\overline{\varepsilon^1_{rv}}$	Random uncertainty associated with first-order moment due to random radial velocity noise	$\mathrm{m\,s^{-1}}$
$\overline{\varepsilon^1_t}$	Random uncertainty associated with first-order moment due to time synchronization between lidar and aircraft	$\mathrm{m\,s^{-1}}$
$\overline{\varepsilon^1_{v_t}}$	Random uncertainty associated with first-order moment due to terminal fall velocity of scattering particles	$\mathrm{m\,s^{-1}}$

Symbol	Description	Unit
$\overline{\beta_{HC}^1}$	Systematic offset uncertainty associated with first-order moment due to contribution of horizontal wind to measured vertical wind, e.g. due to aircraft pitch and roll	$\mathrm{m\,s^{-1}}$
$\overline{\beta_{ac}^1}$	Systematic offset uncertainty associated with first-order moment due to aircraft motion correction	$\mathrm{m\,s^{-1}}$
$\overline{\beta_{b}^1}$	Systematic offset uncertainty associated with first-order moment due to beam pointing-angle calibration	$\mathrm{m\,s^{-1}}$
$\overline{\beta_{rv}^1}$	Systematic offset uncertainty associated with first-order moment due to random radial velocity noise	$\mathrm{m\,s^{-1}}$
$\overline{\beta_{t}^1}$	Random uncertainty associated with first-order moment due to time synchronization between lidar and aircraft	$\mathrm{m\,s^{-1}}$
$\overline{\beta_{v_t}^1}$	Systematic offset uncertainty associated with first-order moment due to terminal fall velocity of scattering particles	$\mathrm{m\,s^{-1}}$
$\overline{\varepsilon_{HC}^2}$	Random uncertainty associated with second-order moment due to positive or negative contribution of horizontal wind to measured variance, e.g. due to non-zero mean aircraft pitch and roll	$\mathrm{m^2\,s^{-2}}$
$\overline{\varepsilon_{PVA}^2}$	Random uncertainty associated with second-order moment due to estimation of systematic pulse-volume averaging effect $\overline{\beta_{PVA}^2}$	$\mathrm{m^2\,s^{-2}}$
$\overline{\varepsilon_{ran}^2}$	Random uncertainty associated with second-order moment due to substitution of ensemble mean with time series mean	$\mathrm{m^2\,s^{-2}}$
$\overline{\varepsilon_{rv}^2}$	Random uncertainty associated with second-order moment due to estimation of systematic random radial velocity noise offset $\overline{\beta_{rv}^2}$	$\mathrm{m^2\,s^{-2}}$
$\overline{\varepsilon_{sys}^2}$	Random uncertainty associated with second-order moment due to estimation of systematic offset $\overline{\beta_{sys}^2}$	$\mathrm{m^2\,s^{-2}}$
$\overline{\beta_{HC}^2}$	Systematic offset uncertainty associated with second-order moment due to positive contribution of horizontal wind to measured variance, e.g. due to variance in aircraft pitch and roll	$\mathrm{m^2\,s^{-2}}$

Symbol	Description	Unit
$\overline{\beta_{PVA}^2}$	Systematic offset uncertainty associated with second-order moment due to pulse-volume averaging	$m^2\,s^{-2}$
$\overline{\beta_{ac}^2}$	Systematic offset uncertainty associated with second-order moment due to aircraft motion correction	$m^2\,s^{-2}$
$\overline{\beta_b^2}$	Systematic offset uncertainty associated with second-order moment due to beam pointing-angle calibration	$m^2\,s^{-2}$
$\overline{\beta_{rv}^2}$	Systematic offset uncertainty associated with second-order moment due to uncorrelated noise in lidar signal	$m^2\,s^{-2}$
$\overline{\beta_{sys}^2}$	Systematic offset uncertainty associated with second-order moment due to substitution of ensemble mean with time series mean	$m^2\,s^{-2}$
$\overline{\beta_t^2}$	Systematic offset uncertainty associated with second-order moment due to time synchronization between lidar and aircraft	$m^2\,s^{-2}$
$\overline{\beta_{v_t}^2}$	Systematic offset uncertainty associated with second-order moment due to terminal fall velocity of scattering particles	$m^2\,s^{-2}$

Bibliography

Adler, B., N. Kalthoff, and O. Kiseleva, 2020a: Detection of structures in the horizontal wind field over complex terrain using coplanar Doppler lidar scans. *Meteorol. Z.*, doi:10.1127/metz/2020/1031.

Adler, B., O. Kiseleva, N. Kalthoff, and A. Wieser, 2019: Comparison of convective boundary layer characteristics from aircraft and wind lidar observations. *J. Atmos. Ocean. Technol.*, **36 (7)**, 1381–1399, doi:10.1175/jtech-d-18-0118.1.

Adler, B., and Coauthors, 2020b: CROSSINN-a field experiment to study the three-dimensional flow structure in the Inn Valley, Austria. *Bull. Amer. Meteor. Soc.*, 1–55, doi:10.1175/BAMS-D-19-0283.1.

Andreae, M. O., and Coauthors, 2015: The Amazon Tall Tower Observatory (ATTO): overview of pilot measurements on ecosystem ecology, meteorology, trace gases, and aerosols. *Atmos. Chem. Phys.*, **15 (18)**, 10 723–10 776, doi:10.5194/acp-15-10723-2015.

Augere, B., and Coauthors, 2019: Three-dimensional wind measurements with the fibered airborne coherent Doppler wind lidar LIVE. *Atmosphere*, **10 (9)**, 549–559, doi:10.3390/atmos10090549.

Baidar, S., S. C. Tucker, M. Beaubien, and R. M. Hardesty, 2018: The optical autocovariance wind lidar. Part II: Green OAWL (GrOAWL) airborne performance and validation. *J. Atmos. Ocean. Technol.*, **35 (10)**, 2099–2116, doi:10.1175/jtech-d-18-0025.1.

Baker, W. E., and Coauthors, 1995: Lidar-measured winds from space: A key component for weather and climate prediction. *Bull. Amer. Meteor. Soc.*, **76 (6)**, 869–888, doi:10.1175/1520-0477(1995)076<0869:lmwfsa>2.0.co;2.

Baker, W. E., and Coauthors, 2014: Lidar-measured wind profiles: The missing link in the global observing system. *Bull. Amer. Meteor. Soc.*, **95 (4)**, 543–564, doi:10.1175/BAMS-D-12-00164.1.

Banakh, V. a., I. N. Smalikho, F. Köpp, and C. Werner, 1995: Representativeness of wind measurements with a cw Doppler lidar in the atmospheric boundary layer. *Appl. Opt.*, **34 (12)**, 2055–2067, doi:10.1364/AO.34.002055.

Banakh, V. A., and C. Werner, 2005: Computer simulation of coherent Doppler lidar measurement of wind velocity and retrieval of turbulent wind statistics. *Opt. Eng.*, **44 (7)**, 71 205, doi:10.1117/1.1955167.

Barbieri, L., and Coauthors, 2019: Intercomparison of small unmanned aircraft system (suas) measurements for atmospheric science during the lapse-rate campaign. *Sensors*, **19 (9)**, 2179, doi:10.3390/s19092179.

Batchelor, G. K., 1948: *The theory of homogeneous turbulence*. Cambridge University Press, Cambridge, 224 pp., doi:10.2307/3609796.

Beyon, J. Y., G. J. Koch, and M. J. Kavaya, 2014: Offshore wind measurements using Doppler aerosol wind lidar (DAWN) at NASA Langley research center. *SPIE Def.*, 8, doi:10.1117/12.2050364.

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

Bonin, T. A., and Coauthors, 2017: Evaluation of turbulence measurement techniques from a single Doppler lidar. *Atmos. Meas. Tech.*, **10 (8)**, 3021–3039, doi:10.5194/amt-10-3021-2017.

Bousquet, O., J. Delanoae, and S. Bielli, 2016: Evaluation of 3D wind observations inferred from the analysis of airborne and ground-based radars during HyMeX SOP-1. *Q. J. R. Meteorol. Soc.*, **142 (August)**, 86–94, doi:10.1002/qj.2710.

Bousquet, O., and P. Tabary, 2014: Development of a nationwide real-time 3-D wind and reflectivity radar composite in France. *Q. J. R. Meteorol. Soc.*, **140 (679)**, 611–625, doi:10.1002/qj.2163.

Browning, K., and R. Wexler, 1968: The determination of kinematic properties of a wind field using Doppler radar. *J. Appl. Meteorol.*, **7 (1)**, 105–113, doi:10.1175/1520-0450(1968)007<0105:tdokpo>2.0.co;2.

Brugger, P., 2014: Implementierung eines Verfahrens zur Ableitung von Kenngrößen der atmosphärischen Turbulenz aus Doppler-Lidar-Messungen der Strukturfunktion. 1–111 pp.

Brugger, P., K. Träumner, and C. Jung, 2016: Evaluation of a procedure to correct spatial averaging in turbulence statistics from a Doppler lidar by comparing time series with an ultrasonic anemometer. *J. Atmos. Ocean. Technol.*, **33 (10)**, 2135–2144, doi:10.1175/jtech-d-15-0136.1.

Bucci, L. R., C. O'Handley, G. D. Emmitt, J. A. Zhang, K. Ryan, and R. Atlas, 2018: Validation of an airborne Doppler wind lidar in tropical cyclones. *Sensors*, **18 (12)**, 4288, doi:10.3390/s18124288.

Caughey, S. J., and S. G. Palmer, 1979: Some aspects of turbulence structure through the depth of the convective boundary layer. *Q. J. R. Meteorol. Soc.*, **105 (446)**, 811–827, doi:10.1002/qj.49710544606.

Cheong, B. L., T. Y. Yu, R. D. Palmer, K. F. Yang, M. W. Hoffman, S. J. Frasier, and F. J. Lopez-Dekker, 2008: Effects of wind field inhomogeneities on Doppler beam swinging revealed by an imaging radar. *J. Atmos. Ocean. Technol.*, **25 (8)**, 1414–1422, doi:10.1175/2007JTECHA969.1.

Chouza, F., O. Reitebuch, A. Benedetti, and B. Weinzierl, 2016a: Saharan dust long-range transport across the Atlantic studied by an airborne Doppler wind lidar and the MACC model. *Atmos. Chem. Phys.*, **16 (18)**, 11 581–11 600, doi:10.5194/acp-16-11581-2016.

Chouza, F., O. Reitebuch, S. Groß, S. Rahm, V. Freudenthaler, C. Toledano, and B. Weinzierl, 2015: Retrieval of aerosol backscatter and extinction from airborne coherent Doppler wind lidar measurements. *Atmos. Meas. Tech.*, **8 (7)**, 2909–2926, doi:10.5194/amt-8-2909-2015.

Chouza, F., O. Reitebuch, M. Jähn, S. Rahm, and B. Weinzierl, 2016b: 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 (7)**, 4675–4692, doi:10.5194/acp-16-4675-2016.

Cornman, L. B., C. S. Morse, and G. Cunning, 1995: Real-time estimation of atmospheric turbulence severity from in-situ aircraft measurements. *J. Aircr.*, **32 (1)**, 171–177, doi:10.2514/3.46697.

Corsmeier, U., R. Hankers, and A. Wieser, 2001: Airborne turbulence measurements in the lower troposphere onboard the research aircraft Dornier 128-6, D-IBUF. *Meteorol. Z.*, **10 (4)**, 315–329, doi:10.1127/0941-2948/2001/0010-0315.

Damiani, R., and S. Haimov, 2006: A high-resolution dual-Doppler technique for fixed multiantenna airborne radar. *IEEE Trans. Geosci. Remote Sens.*, **44 (12)**, 3475–3489, doi: 10.1109/TGRS.2006.881745.

Damiani, R., S. Haimov, and G. Vali, 2005: High resolution airborne radar dual Doppler technique. *11th Conf. Mesoscale Process. 32nd Conf. Radar Meteorol.*

Damiani, R., and Coauthors, 2008: The cumulus, photogrammetric, in situ, and Doppler observations experiment of 2006. *Bull. Amer. Meteor. Soc.*, **89 (1)**, 57–74, doi:10.1175/bams-89-1-57.

Davies, F., C. G. Collier, G. N. Pearson, and K. E. Bozier, 2004: Doppler lidar measurements of turbulent structure function over an urban area. *J. Atmos. Ocean. Technol.*, **21 (5)**, 753–761, doi:10.1175/1520-0426(2004)021<0753:DLMOTS>2.0.CO;2.

De Wekker, S. F. J., K. S. Godwin, G. D. Emmitt, and S. Greco, 2012: Airborne Doppler lidar measurements of valley flows in complex coastal terrain. *J. Appl. Meteorol. Climatol.*, **51 (8)**, 1558–1574, doi:10.1175/JAMC-D-10-05034.1.

De Wekker, S. F. J., and M. Kossmann, 2015: Convective boundary layer heights over mountainous terrain—a review of concepts. *Front. Earth Sci.*, **3**, 22, doi:10.3389/feart.2015.00077.

Deardorff, W. J., 1980: Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-Layer Meteorol.*, **18**, 495–527, doi:10.1007/bf00119502.

Demtröder, W., 2006: *Experimentalphysik 3: Atome, Moleküle und Festkörper*. 3rd ed., Springer Berlin Heidelberg, Berlin, 617 pp.

Deutscher Wetterdienst, 2020a: Das aktuelle numerische Wettervorhersage- und Notfallsystem. URL https://www.dwd.de/DE/forschung/wettervorhersage/num_modellierung/06_num_wettervorhersage_notfallsystem/num_wettervorhersage_notfallsystem_node.html, Website last accessed 27/09/2020.

Deutscher Wetterdienst, 2020b: Datenassimilation. URL https://www.dwd.de/DE/forschung/wettervorhersage/num_modellierung/02_datenassimilation/datenassimilation_node.html, Website last accessed 27/09/2020.

Deutscher Wetterdienst, 2020c: Satellitendaten helfen Problem fehlender Flugzeugmessungen abzumildern. URL https://www.dwd.de/DE/presse/pressemitteilungen/DE/2020/20200519_aeolus_news.html, Website last accessed 27/09/2020.

Didlake, A. C., G. M. Heymsfield, L. Tian, and S. R. Guimond, 2015: The coplane analysis technique for three-dimensional wind retrieval using the HIWRAP Airborne Doppler Radar. *J. Appl. Meteorol. Climatol.*, **54 (3)**, 605–623, doi:10.1175/JAMC-D-14-0203.1.

Dorst, N. M., 2007: The national hurricane research project: 50 years of research, rough rides, and name changes. *Bull. Amer. Meteor. Soc.*, **88 (10)**, 1566–1588, doi:10.1175/BAMS-88-10-1566.

Eberhard, W. L., R. E. Cupp, and K. R. Healy, 1989: Doppler lidar measurement of profiles of turbulence and momentum flux. *J. Atmos. Oceanic Technol.*, **6 (5)**, 809–819, doi:10.1175/1520-0426(1989)006<0809:DLMOPO>2.0.CO;2.

Ellis, S. M., P. Tsai, C. Burghart, U. Romatschke, M. Dixon, J. Vivekanandan, J. Emmett, and E. Loew, 2019: Use of the Earth's surface as a reference to correct airborne nadir-looking radar radial velocity measurements for platform motion. *J. Atmos. Ocean. Technol.*, **36 (8)**, 1343–1360, doi:10.1175/JTECH-D-19-0019.1.

Emeis, S., N. Kalthoff, B. Adler, E. Pardyjak, A. Paci, and W. Junkermann, 2018: High-resolution observations of transport and exchange processes in mountainous terrain. *Atmosphere*, **9 (12)**, 457–474, doi:10.3390/atmos9120457.

Etling, D., 2008: *Theoretische Meteorologie: Eine Einführung*. 3rd ed., Springer Berlin Heidelberg, Berlin, Heidelberg, 376 pp.

European Space Agency, 2018: ESA's Aeolus wind satellite launched. URL https://www.esa.int/For_Media/Press_Releases/ESA_s_Aeolus_wind_satellite_launched, 1 pp.

Eymard, L., and A. Weill, 1988: Dual Doppler radar investigation of the tropical convective boundary layer. *J. Atmos. Sci.*, **45 (5)**, 853–864, doi:10.1175/1520-0469(1988)045<0853:ddriot>2.0.co;2.

Featherstone, W., and S. Claessens, 2007: Closed-form transformation between geodetic and ellipsoidal coordinates. *Stud. Geophys. Geod.*, 1–18, doi:10.1007/s11200-008-0002-6.

Fernando, H. J. S., and Coauthors, 2015: The MATERHORN: Unraveling the intricacies of mountain weather. *Bull. Amer. Meteor. Soc.*, **96**, 1945–1967, doi:10.1175/bams-d-13-00131.1.

Fernando, H. J. S., and Coauthors, 2019: The Perdigão: Peering into microscale details of mountain winds. *Bull. Amer. Meteor. Soc.*, **100**, 799–819, doi:10.1175/bams-d-17-0227.1.

Fiedler, F., and H. A. Panofsky, 1970: Atmospheric scales and spectral gaps. *Bull. Amer. Meteor. Soc.*, **51 (12)**, 1114–1120, doi:10.1175/1520-0477(1970)051<1114:asasg>2.0.co;2.

Frehlich, R., 1997: Effects of wind turbulence on coherent Doppler lidar performance. *J. Atmos. Ocean. Technol.*, **14 (1)**, 54–75, doi:10.1175/1520-0426(1997)014<0054:EOWTOC>2.0.CO;2.

Frehlich, R., 2000: Simulation of coherent Doppler lidar performance for space-based platforms. *J. Appl. Meteorol.*, **39**, 245–262, doi:10.1175/1520-0450(2000)039<0245:socdlp>2.0.co;2.

Frehlich, R., 2001: Errors for space-based Doppler lidar wind measurements: Definition, performance, and verification. *J. Atmos. Ocean. Technol.*, **18 (11)**, 1749–1772, doi:10.1175/1520-0426(2001)018<1749:EFSBDL>2.0.CO;2.

Frehlich, R., and L. Cornman, 2002: Estimating spatial velocity statistics with coherent Doppler lidar. *J. Atmos. Ocean. Technol.*, **19 (3)**, 355–366, doi:10.1175/1520-0426-19.3.355.

Frehlich, R., S. M. Hannon, and S. W. Henderson, 1994: Performance of a 2-μm coherent Doppler lidar for wind measurements. *J. Atmos. Ocean. Technol.*, **11 (6)**, 1517–1528, doi:10.1175/1520-0426(1994)011<1517:poacdl>2.0.co;2.

Frehlich, R., S. M. Hannon, and S. W. Henderson, 1998: Coherent Doppler lidar measurements of wind field statistics. *Boundary-Layer Meteorol.*, **86 (2)**, 233–256, doi:10.1023/A:1000676021745.

Frehlich, R., Y. Meillier, M. L. Jensen, B. Balsley, and R. Sharman, 2006: Measurements of boundary layer profiles in an urban environment. *J. Appl. Meteorol. Climatol.*, **45 (6)**, 821–837, doi:10.1175/JAM2368.1.

Gamache, J. F., 2005: Real-time dissemination of hurricane wind fields determined from airborne Doppler radar. Tech. Rep. 1, Hurricane Research Division (HRD), Miami, FL, 38 pp.

Gamache, J. F., F. D. Marks, and F. Roux, 1995: Comparison of three airborne Doppler sampling techniques with airborne in situ wind observations in Hurricane Gustav (1990). *J. Atmos. Ocean. Technol.*, **12 (1)**, 171–181, doi:10.1175/1520-0426(1995)012<0171:COTADS>2.0.CO;2.

Gasch, P., D. Rieger, C. Walter, P. Khain, Y. Levi, P. Knippertz, and B. Vogel, 2017: Revealing the meteorological drivers of the September 2015 severe dust event in the Eastern Mediterranean. *Atmos. Chem. Phys.*, **17 (22)**, 13 573–13 604, doi:10.5194/acp-17-13573-2017.

Gasch, P., A. Wieser, J. K. Lundquist, and N. Kalthoff, 2020: An LES-based airborne Doppler lidar simulator and its application to wind profiling in inhomogeneous flow conditions. *Atmos. Meas. Tech.*, **13 (3)**, 1609–1631, doi:10.5194/amt-13-1609-2020.

Geerts, B., Q. Miao, and Y. Yang, 2011: Boundary layer turbulence and orographic precipitation growth in cold clouds: Evidence from profiling airborne radar data. *J. Atmos. Sci.*, **68 (10)**, 2344–2365, doi:10.1175/JAS-D-10-05009.1.

Geerts, B., and Coauthors, 2018: Recommendations for in situ and remote sensing capabilities in atmospheric convection and turbulence. *Bull. Amer. Meteor. Soc.*, BAMS–D–17–0310.1, doi:10.1175/BAMS-D-17-0310.1.

Godwin, K. S., S. F. J. De Wekker, and G. D. Emmitt, 2012: Retrieving winds in the surface layer over land using an airborne Doppler lidar. *J. Atmos. Ocean. Technol.*, **29 (4)**, 487–499, doi:10.1175/JTECH-D-11-00139.1.

Greco, S., G. D. Emmitt, M. Garstang, and M. Kavaya, 2020: Doppler Aerosol WiNd (DAWN) lidar during CPEX 2017: Instrument performance and data utility. *Remote Sens.*, **12**, 2951, doi:10.3390/rs12182951.

Grund, C. J., R. M. Banta, J. L. George, J. N. Howell, M. J. Post, R. A. Richer, and A. M. Weickmann, 2001: High-resolution Doppler lidar for boundary layer and cloud research. *J. Atmos. Ocean. Technol.*, **18 (3)**, 376–393, doi:10.1175/1520-0426(2001)018<0376:HRDLFB>2.0.CO;2.

Grunwald, J., N. Kalthoff, U. Corsmeier, and F. Fiedler, 1996: Comparison of areally averaged turbulent fluxes over non-homogeneous terrain: results from the EFEDA-field experiment. *Boundary Layer Meteorol.*, **77 (2)**, 105–134, doi:10.1007/bf00119574.

Guimond, S. R., L. Tian, G. M. Heymsfield, and S. J. Frasier, 2014: Wind retrieval algorithms for the IWRAP and HIWRAP airborne Doppler radars with applications to hurricanes. *J. Atmos. Ocean. Technol.*, **31 (6)**, 1189–1215, doi:10.1175/JTECH-D-13-00140.1.

Haimov, S., J. French, B. Geerts, Z. Wang, M. Deng, A. Rodi, and A. Pazmany, 2018: Compact airborne Ka-Band Radar: a new addition to the University of Wyoming aircraft for atmo-

spheric research. *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 897–900.

Haimov, S., and A. Rodi, 2013: Fixed-antenna pointing-angle calibration of airborne Doppler cloud radar. *J. Atmos. Ocean. Technol.*, **30 (10)**, 2320–2335, doi:10.1175/JTECH-D-12-00262.1.

Hasel, M., 2006: Strukturmerkmale und Modelldarstellung der Konvektion über Mittelgebirgen. Ph.D. thesis, Universität Karlsruhe (TH), Karlsruhe, 170 pp.

Hasel, M., C. Kottmeier, U. Corsmeier, and A. Wieser, 2005: Airborne measurements of turbulent trace gas fluxes and analysis of eddy structure in the convective boundary layer over complex terrain. *Atmos. Res.*, **74 (1-4)**, 381–402, doi:10.1016/j.atmosres.2004.06.010.

Hildebrand, P. H., and Coauthors, 1996: The ELDORA/ASTRAIA airborne Doppler weather radar: High-resolution observations from TOGA COARE. *Bull. Amer. Meteor. Soc.*, **77 (2)**, 213–232, doi:10.1175/1520-0477(1996)077<0213:teadwr>2.0.co;2.

Hinze, J. O., 1959: *Turbulence. An Introduction to its Mechanism and Theory.* McGraw-Hill, New York, NY, 586 pp.

Hofmann-Wellenhof, B., H. Lichtenegger, and J. Collins, 2001: *Global Positioning System - Theory and Practice*. 5th ed., Springer-Verlag Wien, 362 pp.

Hogan, R. J., A. L. Grant, A. J. Illingworth, G. N. Pearson, and E. J. O'Connor, 2009: Vertical velocity variance and skewness in clear and cloud-topped boundary layers as revealed by Doppler lidar. *Quart. J. Roy. Meteor. Soc.*, **135 (640)**, 635–643, doi:10.1002/qj.413.

Holleman, I., 2005: Quality control and verification of weather radar wind profiles. *J. Atmos. Ocean. Technol.*, **22 (10)**, 1541–1550, doi:10.1175/JTECH1781.1.

Holmlund, K., 1998: The utilization of statistical properties of satellite-derived atmospheric motion vectors to derive quality indicators. *Weather Forecast.*, **13 (4)**, 1093–1104, doi:10.1175/1520-0434(1998)013<1093:tuospo>2.0.co;2.

Houbolt, J. C., R. Steiner, and K. G. Pratt, 1964: Dynamic response of airplanes to atmospheric turbulence including flight data on input and response. Tech. rep., NASA Langley Research Center, Hampton, VA, 1–115 pp.

Ingleby, B., and Coauthors, 2016: Progress toward high-resolution, real-time radiosonde reports. *Bull. Amer. Meteor. Soc.*, **97 (11)**, 2149–2161, doi:10.1175/bams-d-15-00169.1.

ISO 5725-1:1994, 1994: *Accuracy (trueness and precision) of measurement methods and results-Part 1: General principles and definitions*. International Organization for Standardization, 17 pp.

Ivey, M., R. Petty, D. Desilets, J. Verlinde, and R. Ellingson, 2014: Polar research with unmanned aircraft and tethered balloons. Tech. rep., DOE Office of Science Atmospheric Radiation Measurement (ARM) Program, 36 pp.

JCGM 100:2008, 2008: Evaluation of measurement data—guide to the expression of uncertainty in measurement. Tech. rep., Joint Committee for Guides in Metrology, 120 pp. URL https://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf.

Kaimal, J., J. Wyngaard, D. Haugen, O. Coté, Y. Izumi, S. Caughey, and C. Readings, 1976: Turbulence structure in the convective boundary layer. *J. Atmos. Sci.*, **33 (11)**, 2152–2169, doi:10.1175/1520-0469(1976)033<2152:tsitcb>2.0.co;2.

Kaimal, J. C., J. Wyngaard, Y. Izumi, and O. Coté, 1972: Spectral characteristics of surface-layer turbulence. *Quart. J. Roy. Meteor. Soc.*, **98 (417)**, 563–589, doi:10.1002/qj.49709841707.

Kalthoff, N., B. Adler, and I. Bischoff-Gauss, 2020: Spatio-temporal structure of the boundary layer under the impact of mountain waves. *Meteorol. Z.*, doi:10.1127/metz/2020/1033.

Kalthoff, N., H. Binder, M. Kossmann, R. Vögtlin, U. Corsmeier, F. Fiedler, and H. Schlager, 1998: Temporal evolution and spatial variation of the boundary layer over complex terrain. *Atmos. Environ.*, **32 (7)**, 1179–1194, doi:10.1016/s1352-2310(97)00193-3.

Kalthoff, N., and Coauthors, 2013: Dry and moist convection in the boundary layer over the black forest - a combined analysis of in situ and remote sensing data. *Meteorol. Z.*, **22 (4)**, 445–461, doi:10.1127/0941-2948/2013/0417.

Kavaya, M. J., J. Y. Beyon, G. J. Koch, M. Petros, P. J. Petzar, U. N. Singh, B. C. Trieu, and J. Yu, 2014: The Doppler aerosol wind (DAWN) airborne, wind-profiling coherent-detection lidar system: Overview and preliminary flight results. *J. Atmos. Ocean. Technol.*, **31 (4)**, 826–842, doi:10.1175/JTECH-D-12-00274.1.

Kawabata, T., H. Iwai, H. Seko, Y. Shoji, K. Saito, S. Ishii, and K. Mizutani, 2014: Cloud-resolving 4D-Var assimilation of Doppler wind lidar data on a meso-gamma-scale convective system. *Mon. Weather Rev.*, **142 (12)**, 4484–4498, doi:10.1175/mwr-d-13-00362.1.

Kiemle, C., M. Wirth, A. Fix, S. Rahm, U. Corsmeier, and P. Di Girolamo, 2011: Latent heat flux measurements over complex terrain by airborne water vapour and wind lidars. *Q. J. R. Meteorol. Soc.*, **137 (S1)**, 190–203, doi:10.1002/qj.757.

Kiemle, C., and Coauthors, 2007: Latent heat flux profiles from collocated airborne water vapor and wind lidars during IHOP_2002. *J. Atmos. Ocean. Technol.*, **24 (4)**, 627–639, doi:10.1175/JTECH1997.1.

Klaas, T., L. Pauscher, and D. Callies, 2015: Lidar-mast deviations in complex terrain and their simulation using CFD. *Meteorol. Z.*, **24 (6)**, 591–603, doi:10.1127/metz/2015/0637.

Klocke, D., M. Brueck, C. Hohenegger, and B. Stevens, 2017: Rediscovery of the doldrums in storm-resolving simulations over the tropical Atlantic. *Nat. Geosci.*, **10 (12)**, 891–896, doi:10.1038/s41561-017-0005-4.

Koch, G. J., J. Y. Beyon, L. J. Cowen, M. J. Kavaya, and M. S. Grant, 2014: Three-dimensional wind profiling of offshore wind energy areas with airborne Doppler lidar. *J. Appl. Remote Sens.*, **8 (1)**, 083 662, doi:10.1117/1.JRS.8.083662.

Koscielny, A., 1984: An evaluation of the accuracy of some radar wind profiling techniques. *J. Atmos. Ocean. Technol.*, **1**, 309–320, doi:10.1175/1520-0426(1984)001<0309:aeotao>2.0.co;2.

Kraus, H., 2008: *Grundlagen der Grenzschicht-Meteorologie: Einführung in die Physik der Atmosphärischen Grenzschicht und in die Mikrometeorologie*. Springer, Berlin Heidelberg, 211 pp.

Kristensen, L., D. H. Lenschow, P. Kirkegaard, and M. Courtney, 1989: The spectral velocity tensor for homogeneous boundary-layer turbulence. *Boundary-Layer Meteorol.*, **47 (1-4)**, 149–193, doi:10.1007/BF00122327.

Kropfli, R., 1986: Single Doppler radar measurements of turbulence profiles in the convective boundary layer. *J. Atmos. Ocean. Technol.*, **3 (2)**, 305–314, doi:10.1175/1520-0426(1986)003<0305:sdrmot>2.0.co;2.

Kuchera, E. L., and M. D. Parker, 2006: Severe convective wind environments. *Weather Forecast.*, **21 (4)**, 595–612, doi:10.1175/waf931.1.

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

Lehner, M., and M. W. Rotach, 2018: Current challenges in understanding and predicting transport and exchange in the atmosphere over mountainous terrain. *Atmosphere*, **9 (7)**, 276–306, doi:10.3390/atmos9070276.

Leike, I., J. Streicher, C. Werner, V. Banakh, I. Smalikho, W. Wergen, and A. Cress, 2001: Virtual Doppler lidar instrument. *J. Atmos. Ocean. Technol.*, **18 (9)**, 1447–1456, doi:10.1175/1520-0426(2001)018<1447:vdli>2.0.co;2.

Lenschow, 1972: The measurement of air velocity and temperature using the NCAR Buffalo aircraft measuring system. Tech. Rep. June, NCAR, Boulder, 1–48 pp.

Lenschow, D., 1970: Airplane measurements of planetary boundary layer structure. *J. Appl. Meteorol.*, **9 (6)**, 874–884, doi:10.1175/1520-0450(1970)009<0874:amopbl>2.0.co;2.

Lenschow, D. H., 1986: Aircraft measurements in the boundary layer. *Probing the atmospheric boundary layer*, D. H. Lenschow, Ed., American Meteorlogical Society, Boston, chap. 5, 39–55, doi:10.1007/978-1-944970-14-7_5.

Lenschow, D. H., and L. Kristensen, 1984: Uncorrelated noise in turbulence measurements. *J. Atmos. Ocean. Technol.*, **2**, 68–81, doi:10.1175/1520-0426(1985)002<0068:unitm>2.0.co;2.

Lenschow, D. H., P. B. Krummel, and S. T. Siems, 1999: Measuring entrainment, divergence, and vorticity on the mesoscale from aircraft. *J. Atmos. Ocean. Technol.*, **16 (10)**, 1384–1400, doi:10.1175/1520-0426(1999)016<1384:MEDAVO>2.0.CO;2.

Lenschow, D. H., M. Lothon, S. D. Mayor, P. P. Sullivan, and G. Canut, 2012: A comparison of higher-order vertical velocity moments in the convective boundary layer from lidar with in situ measurements and large-eddy simulation. *Boundary-Layer Meteorol.*, **143 (1)**, 107–123, doi:10.1007/s10546-011-9615-3.

Lenschow, D. H., J. Mann, and L. Kristensen, 1994: How long is long enough when measuring fluxes and other turbulence statistics? *J. Atmos. Ocean. Technol.*, **11 (3)**, 661–673, doi:10.1175/1520-0426(1994)011<0661:HLILEW>2.0.CO;2.

Lenschow, D. H., and B. B. Stankov, 1986: Length scales in the convective boundary layer. *J. Atmos. Sci.*, **43**, 1198–1209, doi:10.1175/1520-0469(1986)043<1198:LSITCB>2.0.CO;2.

Lenschow, D. H., and P. L. Stephens, 1980: The role of thermals in the convective boundary layer. *Boundary-Layer Meteorol.*, **19 (4)**, 509–532, doi:10.1007/BF00122351.

Lenschow, D. H., and J. Sun, 2007: The spectral composition of fluxes and variances over land and sea out to the mesoscale. *Boundary-Layer Meteorol.*, **125 (1)**, 63–84, doi:10.1007/s10546-007-9191-8.

Lenschow, D. H., V. Wulfmeyer, and C. Senff, 2000: Measuring second- through fourth-order moments in noisy data. *J. Atmos. Ocean. Technol.*, **17 (10)**, 1330–1347, doi:10.1175/1520-0426(2000)017<1330:MSTFOM>2.0.CO;2.

Lenschow, D. H., J. Wyngaard, and P. W. T., 1980: Mean-field and second-moment budgets in a baroclinic, convective boundary layer. *J. Atmos. Sci.*, **37 (6)**, 1313 – 1326, doi:10.1175/1520-0469(1980)037<1313:mfasmb>2.0.co;2.

Leon, D., and G. Vali, 1998: Retrieval of three-dimensional particle velocity from airborne Doppler radar data. *J. Atmos. Ocean. Technol.*, **15 (4)**, 860–870, doi:10.1175/1520-0426(1998)015<0860:ROTDPV>2.0.CO;2.

Leon, D., G. Vali, and M. Lothon, 2006: Dual-Doppler analysis in a single plane from an airborne platform. *J. Atmos. Ocean. Technol.*, **23 (1)**, 3–22, doi:10.1175/JTECH1820.1.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20 (2)**, 130–141, doi:10.1007/978-0-387-21830-4_2.

Lorsolo, S., J. Gamache, and A. Aksoy, 2013: Evaluation of the hurricane research division Doppler radar analysis software using synthetic data. *J. Atmos. Ocean. Technol.*, **30 (6)**, 1055–1071, doi:10.1175/JTECH-D-12-00161.1.

Lothon, M., D. H. Lenschow, D. Leon, and G. Vali, 2005: Turbulence measurements in marine stratocumulus with airborne Doppler radar. *Q. J. R. Meteorol. Soc.*, **131 (609)**, 2063–2080, doi:10.1256/qj.04.131.

Lothon, M., D. H. Lenschow, and S. D. Mayor, 2006: Coherence and scale of vertical velocity in the convective boundary layer from a Doppler lidar. *Boundary-Layer Meteorol.*, **121 (3)**, 521–536, doi:10.1007/s10546-006-9077-1.

Lothon, M., D. H. Lenschow, and S. D. Mayor, 2009: Doppler lidar measurements of vertical velocity spectra in the convective planetary boundary layer. *Boundary-Layer Meteorol.*, **132 (2)**, 205–226, doi:10.1007/s10546-009-9398-y.

Lundquist, J. K., M. J. Churchfield, S. Lee, and A. Clifton, 2015: Quantifying error of lidar and sodar Doppler beam swinging measurements of wind turbine wakes using computational fluid dynamics. *Atmos. Meas. Tech.*, **8 (2)**, 907–920, doi:10.5194/amt-8-907-2015.

Lussier, L. L., and Coauthors, 2018: NCAR/EOL airborne phased array radar (APAR): recent progress and future development efforts. *98th American Meteorological Society Annual Meeting*, AMS.

Lux, O., C. Lemmerz, F. Weiler, U. Marksteiner, B. Witschas, S. Rahm, A. Geiß, and O. Reitebuch, 2020: Intercomparison of wind observations from the European space agency's Aeolus satellite mission and the ALADIN airborne demonstrator. *Atmos. Meas. Tech.*, **13**, 2075–2097, doi:10.5194/amt-13-2075-2020.

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

Mahrt, L., 2000: Surface heterogeneity and vertical structure of the boundary layer. *Boundary-Layer Meteorol.*, **96 (1-2)**, 33–62, doi:10.1023/a:1002482332477.

Maronga, B., and Coauthors, 2015: The parallelized large-eddy simulation model (PALM) version 4.0 for atmospheric and oceanic flows: model formulation, recent developments, and future perspectives. *Geosci. Model Dev.*, **8 (8)**, 2515–2551, doi:10.5194/gmd-8-2515-2015.

Maronga, B., and Coauthors, 2020: Overview of the PALM model system 6.0. *Geosci. Model Dev.*, **13 (3)**, 1335–1372, doi:10.5194/gmd-13-1335-2020.

Mauder, M., and Coauthors, 2020: Comparison of turbulence measurements by a CSAT3B sonic anemometer and a high-resolution bistatic Doppler lidar. *Atmos. Meas. Tech.*, **13 (2)**, 969–983, doi:10.5194/amt-13-969-2020.

Maurer, V., N. Kalthoff, A. Wieser, M. Kohler, M. Mauder, and L. Gantner, 2016: Observed spatiotemporal variability of boundary-layer turbulence over flat, heterogeneous terrain. *Atmos. Chem. Phys.*, **16 (3)**, 1377, doi:10.5194/acp-16-1377-2016.

Mayor, S. D., D. H. Lenschow, R. L. Schwiesow, J. Mann, C. L. Frush, and M. K. Simon, 1997: Validation of NCAR 10.6-μm CO2 Doppler lidar radial velocity measurements and comparison with a 915-MHz profiler. *J. Atmos. Ocean. Technol.*, **14 (5)**, 1110–1126, doi:10.1175/1520-0426(1997)014<1110:VONMCD>2.0.CO;2.

Melhauser, C., and F. Zhang, 2016: Development and application of a simplified coplane wind retrieval algorithm using dual-beam airborne Doppler radar observations for tropical cyclone prediction. *Mon. Weather Rev.*, **144 (7)**, 2645–2666, doi:10.1175/MWR-D-15-0323.1.

Menke, W., 2012: *Describing inverse problems*. Elsevier/Academic Press, Oxford UK, 263 pp., doi:10.1016/B978-0-12-397160-9.00001-1.

Middleton, W., and A. Spilhaus, 1953: *Meteorological instruments*. 3rd ed., University of Toronto Press, Toronto, 286 pp., doi:10.3138/9781487572013.

Moninger, W. R., R. D. Mamrosh, and P. M. Pauley, 2003: Automated meteorological reports from commercial aircraft. *Bull. Amer. Meteor. Soc.*, **84 (2)**, 203–216, doi:10.1175/bams-84-2-203.

Muschinski, A., P. P. Sullivan, D. B. Wuertz, R. J. Hill, S. A. Cohn, D. H. Lenschow, and R. J. Doviak, 1999: First synthesis of wind-profiler signals on the basis of large-eddy simulation data. *Radio Sci.*, **34 (6)**, 1437–1459, doi:10.1029/1999RS900090.

Newman, J. F., P. M. Klein, S. Wharton, A. Sathe, T. A. Bonin, P. B. Chilson, and A. Muschinski, 2016: Evaluation of three lidar scanning strategies for turbulence measurements. *Atmos. Meas. Tech.*, **9 (5)**, 1993–2013, doi:10.5194/amt-9-1993-2016.

Newsom, R. K., W. Alan Brewer, J. M. Wilczak, D. E. Wolfe, S. P. Oncley, and J. K. Lundquist, 2017: Validating precision estimates in horizontal wind measurements from a Doppler lidar. *Atmos. Meas. Tech.*, **10 (3)**, 1229–1240, doi:10.5194/amt-10-1229-2017.

NIMA, 2000: World Geodetic System 1984. NIMA Technical Report 8350.2, 3rd Edition, Amendment 1. Tech. rep., National Imagery and Mapping Agency, Department of Defense, Bethesda, Maryland, USA, 175 pp.

Paffrath, U., C. Lemmerz, O. Reitebuch, B. Witschas, I. Nikolaus, and V. Freudenthaler, 2009: 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 (12)**, 2516–2530, doi:10.1175/2009JTECHA1314.1.

Pal, S., S. F. J. de Wekker, and G. D. Emmitt, 2016: Investigation of the spatial variability of the convective boundary layer heights over an isolated mountain: Cases from the MATERHORN-2012 experiment. *J. Appl. Meteorol. Climatol.*, **55 (9)**, 1927–1952, doi:10.1175/JAMC-D-15-0277.1.

Päschke, E., R. Leinweber, and V. Lehmann, 2015: 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 (6)**, 2251–2266, doi:10.5194/amt-8-2251-2015.

Pauscher, L., and Coauthors, 2016: An inter-comparison study of multi- and DBS lidar measurements in complex terrain. *Remote Sens.*, **8 (9)**, 22, doi:10.3390/rs8090782.

Petty, G. W., 2020: Sampling error in aircraft flux measurements based on a high-resolution large eddy simulation of the marine boundary layer. *Atmos. Meas. Tech. Discuss.*, 1–33, doi:10.1002/essoar.10503723.1.

Platis, A., and Coauthors, 2018: First in situ evidence of wakes in the far field behind offshore wind farms. *Sci. Rep.*, **8 (1)**, 1–10, doi:10.1038/s41598-018-20389-y.

Powell, M. D., S. H. Houston, L. R. Amat, and N. Morisseau-Leroy, 1998: The HRD real-time hurricane wind analysis system. *J. Wind Eng. Ind. Aerodyn.*, **77**, 53–64, doi:10.1016/s0167-6105(98)00131-7.

Pu, Z., L. Zhang, and G. D. Emmitt, 2010: Impact of airborne Doppler wind lidar profiles on numerical simulations of a tropical cyclone. *Geophys. Res. Lett.*, **37 (5)**, 1–5, doi:10.1029/2009GL041765.

Raasch, S., and D. Etling, 1991: Numerical simulation of rotating turbulent thermal convection. *Beitr. Phys. Atmos.*, **64**, 185–1999.

Raasch, S., and M. Schröter, 2001: PALM - A large-eddy simulation model performing on massively parallel computers. *Meteorol. Z.*, **10 (5)**, 363–372, doi:10.1127/0941-2948/2001/0010-0363.

Rahm, S., I. Smalikho, and F. Köpp, 2007: Characterization of aircraft wake vortices by airborne coherent Doppler lidar. *J. Aircr.*, **44 (3)**, 799–805, doi:10.2514/1.24401.

Reitebuch, O., C. Lemmerz, O. Lux, U. Marksteiner, B. Witschas, and R. R. Neely, 2017: WindVal-final report FR-Joint DLR-ESA-NASA wind validation for Aeolus. Tech. rep., DLR. OP-PA, 146 pp.

Reitebuch, O., C. Werner, I. Leike, P. Delville, A. C. P. H. Flamant, and D. Engelbart, 2001: Experimental validation of wind profiling performed by the airborne 10 μ m–heterodyne Doppler lidar WIND. *J. Atmos. Ocean. Technol.*, **18**, 1331–1344, doi:10.1175/1520-0426(2001)018<1331:evowpp>2.0.co;2.

Rodríguez, E., C. S. Morris, and J. E. Belz, 2006: A global assessment of the SRTM performance. *Photogramm. Eng. Remote Sensing*, **72 (3)**, 249–260, doi:10.14358/PERS.72.3.249.

Röhner, L., and K. Träumner, 2013: Aspects of convective boundary layer turbulence measured by a dual-Doppler lidar system. *J. Atmos. Ocean. Technol.*, **30 (9)**, 2132–2142, doi:10.1175/jtech-d-12-00193.1.

Rotach, M. W., and Coauthors, 2020: *Multi-scale transport and exchange processes in the atmosphere over mountains. Programme and experiment*. Innsbruck University Press, Innsbruck, 40 pp., doi:10.15203/99106-003-1.

Sathe, A., and J. Mann, 2013: A review of turbulence measurements using ground-based wind lidars. *Atmos. Meas. Tech.*, **6**, 3147–3167, doi:10.5194/amt-6-3147-2013.

Sawada, M., T. Sakai, T. Iwasaki, H. Seko, K. Saito, and T. Miyoshi, 2015: Assimilating high-resolution winds from a Doppler lidar using an ensemble Kalman filter with lateral boundary adjustment. *Tellus A*, **67 (1)**, 23 473, doi:10.3402/tellusa.v67.23473.

Schäfler, A., B. Harvey, J. Methven, J. D. Doyle, S. Rahm, O. Reitebuch, F. Weiler, and B. Witschas, 2020: Observation of jet stream winds during NAWDEX and characterization of systematic meteorological analysis errors. *Mon. Weather Rev.*, **148 (7)**, 2889–2907, doi:10.1175/mwr-d-19-0229.1.

Schroeder, P., W. A. Brewer, A. Choukulkar, A. Weickmann, M. Zucker, M. W. Holloway, and S. Sandberg, 2020: A compact, flexible, and robust micropulsed Doppler lidar. *J. Atmos. Ocean. Technol.*, **37**, 1387–1402, doi:10.1175/jtech-d-19-0142.1.

Schröter, M., J. Bange, and S. Raasch, 2000: Simulated airborne flux measurements in a LES generated convective boundary layer. *Boundary-Layer Meteorol.*, **95 (3)**, 437–456, doi:10.1023/a:1002649322001.

Scipion, D., 2011: Characterization of the convective boundary layer through a combination of large-eddy simulations and a radar simulator. Phd thesis, University of Oklahoma, 140 pp.

Scipión, D., R. Palmer, P. Chilson, E. Fedorovich, and A. Botnick, 2009: Retrieval of convective boundary layer wind field statistics from radar profiler measurements in conjunction with large eddy simulation. *Meteorol. Z.*, **18 (2)**, 175–187, doi:10.1127/0941-2948/2009/0371.

Scipión, D. E., P. B. Chilson, E. Fedorovich, and R. D. Palmer, 2008: Evaluation of an LES-based wind profiler simulator for observations of a daytime atmospheric convective boundary layer. *J. Atmos. Ocean. Technol.*, **25 (8)**, 1423–1436, doi:10.1175/2007JTECHA970.1.

Seinfeld, J. H., and S. N. Pandis, 2006: *Atmospheric chemistry and physics: from air pollution to climate change*. 2nd ed., John Wiley & Sons Inc, Hoboken, NJ, USA, 1232 pp., doi: 10.1063/1.882420.

Sharman, R. D., L. B. Cornman, G. Meymaris, J. Pearson, and T. Farrar, 2014: Description and derived climatologies of automated in situ eddy-dissipation-rate reports of atmospheric turbulence. *J. Appl. Meteorol. Climatol.*, **53 (6)**, 1416–1432, doi:10.1175/JAMC-D-13-0329.1.

Shenghui, Z., W. Ming, W. Lijun, Z. Chang, and Z. Mingxu, 2014: Sensitivity analysis of the VVP wind retrieval method for single-Doppler weather radars. *J. Atmos. Ocean. Technol.*, **31 (6)**, 1289–1300, doi:10.1175/JTECH-D-13-00190.1.

Stawiarski, C., 2014: *Optimizing dual-Doppler lidar measurements of surface layer coherent structures with large-eddy simulations*. KIT Scientific Publishing, Karlsruhe, 233 pp.

Stawiarski, C., K. Traumner, C. Knigge, and R. Calhoun, 2013: Scopes and challenges of dual-Doppler lidar wind measurements - an error analysis. *J. Atmos. Ocean. Technol.*, **30 (9)**, 2044–2062, doi:10.1175/JTECH-D-12-00244.1.

Stawiarski, C., K. Träumner, C. Kottmeier, C. Knigge, and S. Raasch, 2015: Assessment of surface-layer coherent structure detection in dual-Doppler lidar data based on virtual measurements. *Boundary-Layer Meteorol.*, **156 (3)**, 371–393, doi:10.1007/s10546-015-0039-3.

Storch, H. V., and F. W. Zwiers, 1999: *Statistical analysis in climate research*. Cambridge University Press, Cambridge, UK, 1375 pp., doi:10.1017/CBO9780511612336.

Strauss, L., S. Serafin, S. Haimov, and V. Grubišić, 2015: Turbulence in breaking mountain waves and atmospheric rotors estimated from airborne in situ and Doppler radar measurements. *Q. J. R. Meteorol. Soc.*, **141 (693)**, 3207–3225, doi:10.1002/qj.2604.

Stull, R. B., 1988: *An introduction to boundary layer meteorology*. 13th ed., Springer Science & Business Media, Amsterdam, 681 pp.

Sühring, M., S. Metzger, K. Xu, D. Durden, and A. Desai, 2019: Trade-offs in flux disaggregation: A large-eddy simulation study. *Boundary-Layer Meteorol.*, **170 (1)**, 69–93, doi:10.1007/s10546-018-0387-x.

Sühring, M., and S. Raasch, 2013: Heterogeneity-induced heat-flux patterns in the convective boundary layer: Can they be detected from observations and is there a blending height?—A large-eddy simulation study for the LITFASS-2003 experiment. *Boundary-Layer Meteorol.*, **148 (2)**, 309–331, doi:10.1007/s10546-013-9822-1.

Taylor, G. I., 1938: The spectrum of turbulence. *Proc. R. Soc. London. Ser. A, Math. Phys. Sci.*, **164 (919)**, 476–490, doi:10.1098/rspa.1938.0032.

Thiébaux, H. J., and F. W. Zwiers, 1984: The interpretation and estimation of effective sample size. *J. Appl. Meteorol.*, **23 (5)**, 800–811, doi:10.1175/1520-0450(1984)023<0800:tiaeoe>2.0.co;2.

Tian, L., G. M. Heymsfield, A. C. Didlake, S. Guimond, and L. Li, 2015: Velocity-azimuth display analysis of Doppler velocity for HIWRAP. *J. Appl. Meteorol. Climatol.*, **54 (8)**, 1792–1808, doi:10.1175/JAMC-D-14-0054.1.

Tompkins, A. M., 2001: Organization of tropical convection in low vertical wind shears: The role of cold pools. *Mon. Weather Rev.*, **58**, 1650–1672, doi:10.1175/1520-0469(2001)058<1650:ootcil>2.0.co;2.

Träumner, K., 2010: *Einmischprozesse am Oberrand der konvektiven atmosphärischen Grenzschicht*. KIT Scientific Publishing, Karlsruhe, 245 pp.

Träumner, K., T. Damian, C. Stawiarski, and A. Wieser, 2015: Turbulent structures and coherence in the atmospheric surface layer. *Boundary-Layer Meteorol.*, **154 (1)**, 1–25, doi:10.1007/s10546-014-9967-6.

Tucker, S. C., C. S. Weimer, S. Baidar, and R. M. Hardesty, 2018: The optical autocovariance wind lidar. Part I: OAWL instrument development and demonstration. *J. Atmos. Ocean. Technol.*, **35 (10)**, 2079–2097, doi:10.1175/JTECH-D-18-0024.1.

Turk, F. J., S. Hristova-Veleva, S. L. Durden, S. Tanelli, O. Sy, G. D. Emmitt, S. Greco, and S. Q. Zhang, 2020: Joint analysis of convective structure from the APR-2 precipitation radar and the DAWN Doppler wind lidar during the 2017 convective processes experiment (CPEX). *Atmos. Meas. Tech.*, **13 (8)**, 4521–4537, doi:10.5194/amt-13-4521-2020.

Van Zyl, J. J., 2001: The shuttle radar topography mission (SRTM): A breakthrough in remote sensing of topography. *Acta Astronaut.*, **48 (5-12)**, 559–565, doi:10.1016/S0094-5765(01)00020-0.

Vihma, T., and C. Kottmeier, 2000: A modelling approach for optimizing flight patterns in airborne meteorological measurements. *Boundary Layer Meteorol.*, **95 (2)**, 211–230, doi:10.1023/a:1002634613282.

von Karman, T., 1948: Progress in the statistical theory of turbulence. *Proc. N. A. S.*, **34**, 530–539, doi:10.1073/pnas.34.11.530.

Wainwright, C. E., P. M. Stepanian, P. B. Chilson, R. D. Palmer, E. Fedorovich, and J. A. Gibbs, 2014: A time series sodar simulator based on large-eddy simulation. *J. Atmos. Ocean. Technol.*, **31 (4)**, 876–889, doi:10.1175/JTECH-D-13-00161.1.

Waldteufel, P., and H. Corbin, 1978: On the analysis of single-Doppler radar data. *J. Appl. Meteorol.*, **18**, 532–542, doi:10.1175/1520-0450(1979)018<0532:OTAOSD>2.0.CO;2.

Wandlinger, U., 2005: Introduction to lidar. *Lidar - range resolved optical remote sensing of the atmosphere*, Springer Science & Business Media, New York, NY, chap. 1, 1–18.

Wang, H., R. J. Barthelmie, A. Clifton, and S. C. Pryor, 2015: Wind measurements from arc scans with Doppler wind lidar. *J. Atmos. Ocean. Technol.*, **32 (11)**, 2024–2040, doi:10.1175/JTECH-D-14-00059.1.

Weissmann, M., 2006: Meteorological applications of airborne Doppler lidar measurements. Ph.D. thesis, Leopold-Franzens-Universität Innsbruck, 84 pp.

Weissmann, M., F. Braun, L. Gantner, G. Mayr, S. Rahm, and O. Reitebuch, 2005a: The Alpine Mountain – Plain Circulation : Airborne Doppler Lidar Measurements and. *Mon. Weather Rev.*, **133**, 3095–3109, doi:10.1175/MWR3012.1.

Weissmann, M., R. Busen, A. Dörnbrack, S. Rahm, and O. Reitebuch, 2005b: Targeted observations with an airborne wind lidar. *J. Atmos. Ocean. Technol.*, **22 (11)**, 1706–1719, doi:10.1175/JTECH1801.1.

Weissmann, M., R. H. Langland, C. Cardinali, P. M. Pauley, and S. Rahm, 2012: Influence of airborne Doppler wind lidar profiles near Typhoon Sinlaku on ECMWF and NOGAPS forecasts. *Q. J. R. Meteorol. Soc.*, **138 (662)**, 118–130, doi:10.1002/qj.896.

Weitkamp, C., Ed., 2005: *Lidar - range-resolved optical remote sensing of the atmosphere*. Springer Science & Business Media, New York, NY, 455 pp.

Wendisch, M., and J.-L. Brenguier, 2013: Measuring the three-dimensional wind vector using a five-hole probe. *Airborne Measurements for Environmental Research*, Wiley-VCH, Berlin, chap. A1, 1–9.

Werner, C., 2005: Doppler wind lidar. *Lidar - range-resolved optical remote sensing of the atmosphere*, Springer Science & Business Media, New York, NY, chap. 12, 326–354.

Witschas, B., C. Lemmerz, A. Geiß, O. Lux, U. Marksteiner, S. Rahm, O. Reitebuch, and F. Weiler, 2020: First validation of Aeolus wind observations by airborne Doppler wind lidar measurements. *Atmos. Meas. Tech.*, **13 (5)**, 2381–2396, doi:10.5194/amt-13-2381-2020.

Witschas, B., S. Rahm, A. Dörnbrack, J. Wagner, and M. Rapp, 2017: Airborne wind lidar measurements of vertical and horizontal winds for the investigation of orographically induced gravity waves. *J. Atmos. Ocean. Technol.*, **34 (6)**, 1371–1386, doi:10.1175/JTECH-D-17-0021.1.

Zhang, J., R. Atlas, G. Emmitt, L. Bucci, and K. Ryan, 2018: Airborne Doppler wind lidar observations of the tropical cyclone boundary layer. *Remote Sens.*, **10 (6)**, 825, doi:10.3390/rs10060825.