LiSBOA: LiDAR Statistical Barnes Objective Analysis for optimal design of LiDAR scans and retrieval of wind statistics. Part I: Theoretical framework

Stefano Letizia, Lu Zhan, and Giacomo Valerio Iungo

Wind Fluids and Experiments (WindFluX) Laboratory, Mechanical Engineering Department, The University of Texas at Dallas, 800 W Campbell Rd, 75080 Richardson, TX, USA

Correspondence: Giacomo Valerio Iungo, email: valerio.iungo@utdallas.edu

Abstract. A LiDAR Statistical Barnes Objective Analysis (LiSBOA) for the optimal design of LiDAR scans and retrieval of the velocity statistical moments is proposed. The LiSBOA represents an adaptation of the classical Barnes scheme for the statistical analysis of unstructured experimental data in N-dimensional space and it is a suitable technique for the evaluation over a structured Cartesian grid of the statistics of scalar fields sampled through scanning LiDARs. The LiSBOA is validated

5   and characterized via a Monte Carlo approach applied to a synthetic velocity field. This revisited theoretical framework for the Barnes objective analysis enables the formulation of guidelines for the optimal design of LiDAR experiments and efficient application of the LiSBOA for the post-processing of LiDAR measurements. The optimal design of LiDAR scans is formulated as a two cost-function optimization problem including the minimization of the percentage of the measurement volume not sampled with adequate spatial resolution and the minimization of the error on the mean of the velocity field. The optimal

10   design of the LiDAR scans also guides the selection of the smoothing parameter and the total number of iterations to use for the Barnes scheme. The LiSBOA is assessed against a numerical dataset generated using the virtual LiDAR technique applied to the data obtained from a large eddy simulation (LES). The optimal sampling parameters for a scanning Doppler pulsed wind LiDAR are retrieved through the LiSBOA, then the estimated statistics are compared with those of the original LES dataset, showing a maximum error of about 4% for both mean velocity and turbulence intensity.

15 List of symbols

- x, y, z: streawmise, spanwise, vertical Cartesian coordinates

- t: non-spatial coordinate or time

- u, v, w: streamwise, spanwise, vertical velocity components

- u_{LOS}: radial or line-of-sight wind speed

- L: number of realizations/scans

- θ: azimuth angle

- β: elevation angle

- $\Delta\theta$: azimuth-angle resolution

- $\Delta\beta$: elevation-angle resolution

- τ_a: accumulation time

- Δr: gate length

- N_r: number of range gates along the laser beam

- T: total sampling time

- σ: smoothing parameter

- m: number of iterations

- R_{max}: radius of influence

- $\boldsymbol{\Delta n}$: half-wavelength vector

- $\boldsymbol{\Delta n_0}$: fundamental half-wavelength vector

- Δd: random data spacing

- $\boldsymbol{dx}$: resolution vector in Cartesian coordinates

- D^m: response at the m-th iteration

- $\boldsymbol{r_i}$: position in the N-dimensional space of the i-th grid node of the Cartesian grid

- $\boldsymbol{r_j}$: position in the N-dimensional space of the j-th sample

- ϵ^I: cost function I (data loss)

40 - ϵ^{II}: cost function II (standard deviation of the sample mean)

- N_i: total number of nodes of the Cartesian grid

- τ: integral time-scale

- $\tilde{\cdot}$: spatial variable in the scaled frame of reference

1 Introduction

Reliable measurements of the wind-velocity vector field are essential to understand the complex nature of atmospheric turbulence and provide valuable datasets for the validation of theoretical and numerical models. However, field measurements of wind speed are typically characterized by large uncertainties due to the generally unknown and uncontrollable boundary conditions (Braham, 1979), the broad range of time and length-scales (Cushman-Roisin and Beckers, 1990a), and the complexity of the physics involved (Stull, 1988). Furthermore, the large measurement volume, which typically extends throughout the height of the atmospheric boundary layer, imposes to the experimentalists the selection of the sampling parameters as a trade-off between spatial and temporal resolutions.

Wind speed has been traditionally measured through local sensors, such as mechanical, sonic, and hot-wire anemometers (Liu et al., 2019; Kunkel and Marusic, 2006). Besides their simplicity, mechanical anemometers are affected by errors due to the flow distortion of the supporting structures, drawbacks under harsh weather conditions (Mortensen, 1994), and over-speeding (Busch and Kristensen, 1976). Furthermore, their relatively slow response results in a limited range of the measurable time-length scales, which makes them unsuitable, for instance, to measure the turbulent flow around urban areas (Pardyjak and Stoll, 2017). Sonic anemometers can measure the three velocity component with frequencies up to 100 Hz (Cuerva and Sanz-Andrés, 2000) in a probing volume of the order of 10^{-4} m^3, yet measurements might be still affected by the wakes generated by the supporting structures, such as met-towers and struts and they are sensitive to temperature variations (Mortensen, 1994). Hot-wire anemometers, although they provide a full characterization of the energy spectrum, require a complicated calibration (Kunkel and Marusic, 2006) and are extremely fragile (Wheeler and Ganji, 2010a). Furthermore, traditional single-point sensors are unable to provide an adequate characterization of the spatial gradients of the wind velocity vector, which is particularly significant in the vertical direction (Cushman-Roisin and Beckers, 1990b). To overcome this issue, several anemometers arranged in arrays and supported by meteorological masts have been deployed in several field campaigns (Haugen et al., 1971; Bradley, 1983; Taylor and Teunissen, 1987; Emeis et al., 1995; Pashow et al., 2001; Berg et al., 2011; Kunkel and Marusic, 2006).

In the last few decades, remote sensing instruments have been increasingly utilized to probe the atmospheric boundary layer (Debnath et al., 2017b, a) and nowadays they represent a more cost-effective and flexible alternative to meteorological towers (Newsom et al., 2017). In particular, in the realm of remote sensing anemometry, Doppler wind light detection and ranging (LiDAR) systems underwent a rapid development due to the significant advancement in eye-safe laser technology (Emeis, 2010). Wind LiDARs have been heavily employed in wind energy (Bingöl et al., 2010; Aitken and Lundquist, 2014; Trujillo et al., 2011; Iungo et al., 2013b; Machefaux et al., 2016b; Garcia et al., 2017; El-Asha et al., 2017; Bromm et al., 2018; Zhan et al., 2019, 2020), airport monitoring (Köpp et al., 2005; Tang et al., 2011; Holzäpfel et al., 2016; Thobois et al., 2019), micro-meteorology (Gal-Chen et al., 1992; Banakh et al., 1999; Banta et al., 2006; Mann et al., 2010; Muñoz-Esparza et al., 2012; Rajewski et al., 2013; Schween et al., 2014), urban wind research (Davies et al., 2007; Newsom et al., 2008; Xia et al., 2008; Kongara et al., 2012; Huang et al., 2017; Halios and Barlow, 2018) and studies of terrain-induced effects (Bingöl, 2009; Krishnamurthy et al., 2013; Kim et al., 2016; Pauscher et al., 2016; Risan et al., 2018; Fernando et al., 2019; Bell et al., 2020).

Besides the mentioned capabilities, LiDARs present some important limitations, such as reduced range in adverse weather conditions (precipitations, heavy rain, fog, low clouds or low aerosol concentration) (Liu et al. (2019), Mann et al. (2018)) and a limited spatio-temporal resolution of this instrument, namely about 20 meters in the radial direction and about 10 Hz in sampling frequency. These technical specifications, associated with the non-stationary wind conditions typically encountered for field experiments, pose major challenges to apply wind LiDARs for the statistical analysis of turbulent atmospheric flows.

In the realm of wind energy, early LiDAR measurements were limited to the qualitative analysis of snapshots of the line-of-sight (LOS) velocity, i.e. the velocity component parallel to the laser beam (Käsler et al., 2010; Clive et al., 2011). Fitting of the wake velocity deficit was also successfully exploited to extract quantitative information about wake evolution from LiDAR measurements (Aitken and Lundquist, 2014; Wang and Barthelmie, 2015; Kumer et al., 2015; Trujillo et al., 2016; Bodini et al., 2017). To characterize velocity fields with higher statistical significance, the time averages of several LiDAR scans were calculated for periods with reasonably steady inflow conditions (Iungo and Porté-Agel, 2014; Machefaux et al., 2016b; Van Dooren et al., 2016). In the case of data collected under different wind and atmospheric conditions, clustering and bin-averaging of LiDAR data were carried out (Machefaux et al., 2016b; Garcia et al., 2017; Bromm et al., 2018; Zhan et al., 2019, 2020). Finally, more advanced techniques for first-order statistical analysis, such as variational methods (Xia et al., 2008; Newsom and Banta, 2004), optimal interpolation (Xu and Gong, 2002; Kongara et al., 2012), least-squares methods (Newsom et al., 2008), Navier-Stokes solvers (Astrup et al., 2017; Sekar et al., 2018) were applied for the reconstruction of the velocity vector field from dual-Doppler measurements.

Besides the mean field, the calculation of higher-order statistics from LiDAR data to investigate atmospheric turbulence is still an open problem. In this regard, Eberhard et al. (1989) re-adapted the post-processing of the velocity azimuth display (VAD) scans (Lhermitte, 1969; Wilson, 1970; Kropfli, 1986) to estimate all the components of the Reynolds stress tensor by assuming horizontal homogeneity of the mean flow within the scanning volume, which can be a limiting constraint for measurements in complex terrains (Frisch, 1991; Bingöl, 2009). Range height indicator (RHI) scans were used to detect second-order statistics (Bonin et al., 2017), spectra, skewness, dissipation rate of the velocity field, and even heat flux (Gal-Chen et al., 1992). Recently, in the context of wind radar technology, but readily applicable to LiDARs as well, a promising method for the estimation of the instantaneous turbulence intensity (i.e. the ratio between standard deviation and mean of streamwise velocity) based on the Taylor hypothesis of frozen turbulence was proposed by Duncan et al. (2019). More advanced techniques exploit additional information of turbulence carried by the spectrum of the back-scattered LiDAR signal (Smalikho, 1995). However, this approach requires the availability of LiDAR raw data, which is not generally granted for commercial LiDARs. For a review on turbulence statistical analyses through LiDAR measurements, the reader can refer to Sathe and Mann (2013). Another typical scanning strategy to obtain high-frequency LiDAR data consists in performing scans with fixed elevation and azimuthal angles of the laser beam while maximizing the sampling frequency (Mayor et al., 1997; O'Connor et al., 2010; Vakkari et al., 2015; Frehlich and Cornman, 2002; Debnath et al., 2017b; Choukulkar et al., 2017; Lundquist et al., 2017).

For remote sensing instruments, data are typically collected based on a spherical coordinate system, then interpolated over a Cartesian reference frame oriented with the x-axis in the mean wind-direction. This interpolation can be a source of error (Fuertes Carbajo and Porté-Agel, 2018), especially if a linear interpolation method is used (Garcia et al., 2017; Carbajo Fuertes

et al., 2018; Beck and Kühn, 2017; Astrup et al., 2017). Delaunay triangulation has also been widely adopted for coordinate transformation (Clive et al., 2011; Trujillo et al., 2011; Iungo and Porté-Agel, 2014; Trujillo et al., 2016; Machefaux et al., 2016b), yet with accuracy not quantified in case of non-uniformly distributed data. It is reasonable to weight the influence of the experimental points on their statistics by the distance from the respective grid centroid, such as using uniform (Newsom et al., 2008), hyperbolic (Van Dooren et al., 2016), or Gaussian weights (Newsom et al., 2014; Wang and Barthelmie, 2015; Zhan et al., 2019). The use of distance-based Gaussian weights for the interpolation of scattered data over a Cartesian grid is at the base of the Barnes objective analysis (or Barnes scheme) (Barnes, 1964), which has been systematically used in meteorology, but only sporadically for LiDAR data. It represents an iterative statistical ensemble procedure to reconstruct a scalar field arbitrarily sampled in space and low-pass filtered with a cut-off wavelength that is a function of the parameters of the scheme.

The scope of this work is to define a methodology to post-process scattered data of a turbulent velocity field measured through a scanning Doppler wind LiDAR (or eventually other remote sensing instruments) to calculate mean, standard deviation and even higher-order statistical moments on a Cartesian grid. The proposed methodology, referred to as LiDAR Statistical Barnes Objective Analysis (LiSBOA), represents an adaptation of the classic Barnes scheme to N-dimensional domains enabling applications for non-isotropic scalar fields through a coordinate transformation. A major point of novelty of the LiSBOA is the estimation of wind-velocity variance (and eventually higher-order statistics) from the residual field of the mean, which also provides adequate filtering of dispersive stresses due to data variability not connected with the turbulent motion. A criterion for rejection of statistical data affected by aliasing due to the undersampling of the spatial wavelengths under investigation is formulated. The LiSBOA is assessed against a synthetic scalar field to validate its theoretical response and the formulated error metric. Detailed guidelines for the optimal design of a LiDAR experiment and effective reconstruction of the wind statistics are provided. The effectiveness of the proposed scheme in the identification of the optimal scanning parameters and retrieval of turbulence statistics is quantified using virtual LiDAR data.

It will be shown that the revisited Barnes scheme offers several advantages compared to the above-cited techniques for LiDAR data analysis: i) it allows to explicitly select the cut-off wavenumber to filter out small-scale variability while retaining relevant modes in the flow field; ii) the distance-based weighting function provides smoother fields than linear interpolation or box-averaging, while still being simpler and computationally inexpensive compared to more sophisticated techniques (e.g. optimal interpolation, variational methods); iii) it provides guidance for the optimal design of LiDAR scans to investigate specific wavelengths in the flow. On the other hand, the procedure requires estimates of input parameters for the flow under investigation and the LiDAR system used. In case these parameters cannot be obtained from existing literature or preliminary tests, then a sensitivity study on the variability of the LiSBOA results to the input parameters can be carried out.

The remainder of the manuscript is organized as follows: in Sect. 2, the extension of the Barnes scheme theory to N-dimensional domains and higher-order statistical moments is presented. In Sect. 3, the theoretical response function of the LiSBOA is validated against a synthetic case, while guidelines for proper use of the proposed algorithm and optimal scan design are provided in Sect. 4. In Sect. 5, the accuracy of the LiSBOA is tested using of the virtual LiDAR technique. Challenges in

the application of the methodology to field experimental data are then discussed in Sect. 6. Finally, concluding remarks are provided in Sect. 7.

2   The Barnes Objective Analysis: fundamentals and extension to statistical N-dimensional analysis

150 The Barnes scheme was originally conceived as an iterative algorithm aiming to interpolate a set of sparse data over a Cartesian grid (Barnes, 1964) and it was inspired by the successive correction scheme by Cressman (1959). The first iteration of the algorithm calculates a weighted-space-averaged field, g^0, over a Cartesian grid from the sampled scalar field, f. The mean field is iteratively modified by adding contributions to recover features characterized by shorter wavelengths, which are inevitably damped by the initial averaging process. In this work, we adopt the most classical form of the Barnes scheme as follows:

$$\begin{cases} g_i^0 = \sum_j w_{ij} f_j \\ g_i^m = \sum_j w_{ij}(f_j - \phi(g^{m-1})_j) + g_i^{m-1} \quad \forall \quad m \in \mathbb{N}^+, \end{cases} \tag{1}$$

where g_i^m is the average field at the i-th grid node with coordinates $\boldsymbol{r_i}$ (bold symbols indicate vectorial quantities) for the m-th iteration, f_j is the scalar field sampled at the location $\boldsymbol{r_j}$ and ϕ represents the linear interpolation operator from the Cartesian grid to the sample location. The weights for the sample acquired at the location $\boldsymbol{r_j}$ and for the calculation of the statistics of f at the grid node with coordinates $\boldsymbol{r_i}$, w_{ij}, are defined as:

$$w_{ij} = \frac{e^{-\frac{|\boldsymbol{r_i}-\boldsymbol{r_j}|^2}{2\sigma^2}}}{\sum_j e^{-\frac{|\boldsymbol{r_i}-\boldsymbol{r_j}|^2}{2\sigma^2}}}, \tag{2}$$

where σ is referred to as smoothing parameter and $|.|$ indicates Euclidean norm. For practical reasons, the summations over j are performed over the neighboring points included in a ball with a finite radius $R_{\max}$ (also called the radius of influence) and centered at the i-th grid point. In this work, following Barnes (1964), we select $R_{\max} = 3\sigma$, which encompasses 99.7% and 97% of the volume of the weighting function in 2D and 3D, respectively.

165 In literature, there is a lack of consensus for the selection of the total number of iterations (Barnes, 1964; Achtemeier, 1989; Smith and Leslie, 1984; Seaman, 1989) and the smoothing parameter (Barnes, 1994a; Caracena, 1987; Pauley and Wu, 1990). A reduction of the smoothing parameter, σ, as a function of the iteration, m, was originally proposed by Barnes (1973); however, this approach resulted to be detrimental in terms of noise suppression (Barnes, 1994c).

In the frequency domain, the Barnes objective analysis is tractable as a low-pass filter applied to a scalar field, f, with a 170 response as a function of the spatial wavelength depending on the smoothing parameter, σ, and the number of iterations, m. This feature has been exploited in meteorology to separate small-scale from mesoscale motions (Doswell, 1977; Maddox, 1980; Gomis and Alonso, 1990). The spectral behavior of the Barnes scheme has been traditionally characterized by calculating the so-called continuous response at the m-th iteration, $D^m(\boldsymbol{k})$, with $\boldsymbol{k}$ being the wavenumber vector. $D^m(\boldsymbol{k})$ is defined as the ratio between the amplitude of the Fourier mode $e^{i\boldsymbol{k}\cdot\boldsymbol{x}}$ (with $i = \sqrt{-1}$) for the reconstructed field, g^m, to its amplitude for in input

field, f, in the limit of a continuous distribution of samples and infinite domain. The analytical expression for the continuous response was provided by Barnes (1964) and Pauley and Wu (1990) for 1D and 2D domains, respectively, while in the context of the LiSBOA it is extended to N dimensions to enhance its applicability. Furthermore, besides the spatial variability of f, the temporal coordinate, t, is introduced to determine the response of the statistical moments of f.

We consider a continuous scalar field, $f(\boldsymbol{x},t)$, which is defined over an N-dimensional domain, $\boldsymbol{x}$. It is further assumed that the field f is ergodic in time. In practice, ergodic data can be obtained by selecting samples collected for a temporal window exhibiting stationary boundary conditions or, more generally, through a cluster analysis of discontinuous data (Machefaux et al., 2016a; Bromm et al., 2018; Iungo et al., 2018; Zhan et al., 2019, 2020). By adopting the approach proposed by Pauley and Wu (1990) and by taking advantage of the isotropy of the Gaussian weights (Eq. (2)), we can define the LiSBOA operator at the 0-th iteration as:

$$g^0(\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^N} \int_{\mathbb{R}^N} \left[\frac{1}{t_2-t_1} \int_{t_1}^{t_2} f(\boldsymbol{\xi},t)dt \right] e^{-\frac{|\boldsymbol{x}-\boldsymbol{\xi}|^2}{2\sigma^2}} d\boldsymbol{\xi}, \tag{3}$$

where t_1 and t_2 are initial and final time. The term within the square brackets represents the mean of f over the considered sampling interval $[t_1, t_2]$, which is indicated as $\overline{f}$. Moreover, to reconstruct a generic q-th central statistical moment of the scalar field, f, it is sufficient to apply the LiSBOA operator of Eq. (3) to the fluctuations over $\overline{f}$ to the q-th power:

$$\mu_f^q(\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^N} \int_{\mathbb{R}^N} \left\{ \frac{1}{t_2-t_1} \int_{t_1}^{t_2} \left[f(\boldsymbol{\xi},t) - \overline{f}(\boldsymbol{\xi},t) \right]^q dt \right\} e^{-\frac{|\boldsymbol{x}-\boldsymbol{\xi}|^2}{2\sigma^2}} d\boldsymbol{\xi}. \tag{4}$$

For practical applications, the mean field $\overline{f}$ is generally not known, but it can be approximated by the LiSBOA output, g^m, interpolated at the sample location through the operator ϕ. By comparing Eq. (4) with Eq. (3), it is understandable that the response function of any central moment with order higher than one is equal to that of the 0-th iteration response of the mean, g^0. Indeed, Eq. (3) can be interpreted as the 0-th iteration of the LiSBOA spatial operator (viz. Eq. (3)) applied to the fluctuation field to the q-th power.

By leveraging the convolution theorem, it is possible to calculate the response function of the mean of the 0-th iteration of the LiSBOA in the frequency domain (see Appendix A for more details). This result, combined with the recursive formula of Barnes (1964) for the response at the generic iteration m, provides the spectral response of the LiSBOA for the mean:

$$D^m = \begin{cases} D^0(\boldsymbol{k}) = e^{-\frac{\sigma^2}{2}|\boldsymbol{k}|^2} = e^{-\frac{\sigma^2\pi^2}{2}\left[\sum_{p=1}^N \frac{1}{\Delta n_p^2}\right]} & \text{for } m = 0 \\ D^0 \sum_{p=0}^m (1-D^0)^p & \text{for } m \in \mathbb{N}^+, \end{cases} \tag{5}$$

where $\boldsymbol{\Delta n}$ is the half-wavelength vector associated with $\boldsymbol{k}$. Equation 5 states that, for a given wavenumber (i.e. half-wavelength), the respective amplitude of the interpolated scalar field, g^m, is equal to that of the original scalar field, damped with a function of the smoothing parameter, σ, and the number of iterations, m. This implies that the parameters σ and m should be selected properly to avoid significant damping for wavelengths of interest or dominating the spatial variability of the scalar field under investigation.

For real applications, the actual LiSBOA response function can depart from the above-mentioned theoretical response (Eq. (5)) for the following reasons:

- the convolution integral in Eq. (3) is calculated over a ball of finite radius R_{max};

- f is sampled over a discrete domain and, thus, introducing related limitations, such as the risk of aliasing (Pauley and Wu, 1990);

- the distribution of the sampling points is usually irregular and non-uniform leading to larger errors where a lower sample density is present (Smith et al., 1986; Smith and Leslie, 1984; Buzzi et al., 1991; Barnes, 1994a) or in proximity to the domain boundaries (Achtemeier, 1986);

- an error is introduced by the back-interpolation function, ϕ, from the Cartesian grid, $\boldsymbol{r_i}$, to the location of the samples, $\boldsymbol{r_j}$ (Eq. (1)) (Pauley and Wu, 1990).

Before proceeding with further analysis, it is necessary to address the applicability of the LiSBOA to anisotropic and multi-chromatic scalar fields. Generally, the application of the LiSBOA with an isotropic weighting function is not recommended in case of severe anisotropy of the field and/or the data distribution. At the early stages of objective analysis techniques, the use of an anisotropic weighting function was proved to be beneficial to increase accuracy while highlighting patterns elongated along a specific direction, based on empirical (Endlich and Mancuso, 1968) and theoretical arguments (Sasaki, 1971). Furthermore, the adoption of a directional smoothing parameter, σ_p, where p is a generic direction, allows maximizing the utilization of the data retrieved through inherently anisotropic measurements, such as the line-of-sight fields detected by remote sensing instruments (Askelson et al., 2000; Trapp and Doswell, 2000). With this in mind, we propose a linear scaling of the physical coordinates before the application of the LiSBOA to recover a pseudo-isotropic velocity field. The scaling reads as:

$$\tilde{x}_p = \frac{x_p - x_p^*}{\Delta n_{0,p}}, \tag{6}$$

where $\boldsymbol{x}^*$ is the origin of the scaled reference frame and $\Delta n_{0,p}$ is the scaling factor for the p-th direction. Hereinafter, $\tilde{\ }$ refers to the scaled frame of reference. From a physical standpoint, the scaling is equivalent to the adoption of an anisotropic weighting function, while the re-scaling approach is preferred to ensure generality to the mathematical formulation outlined in this section.

The scaling factor, $\boldsymbol{\Delta n_0}$, is an important parameter in the present framework and is referred to as fundamental half-wavelength, while the associated Fourier mode is denoted as the fundamental mode. The selection of the fundamental half-wavelength should be guided by a priori knowledge of the dominant length-scales of the flow in various directions. Modes exhibiting degrees of anisotropy different than that of the selected fundamental mode, will not be isotropic in the scaled mapping, which leads to two consequences: first, their response will not be optimal, in the sense that the shortest directional wavelength can produce excessive damping of the specific mode (Askelson et al., 2000); second, the shape preservation of such non-spherical features in the field reconstructed through the LiSBOA is not ensured (Trapp and Doswell, 2000).

Regarding the reconstruction of the flow statistics through the LiSBOA, two categories of error can be identified. The first is the statistical error due to the finite number of samples of the scalar field, f, available in time. This error is strictly connected with the local turbulence statistics, the sampling rate, and the duration of the experiment. The second error category is the spatial sampling error, which is due to the discrete sampling of f in the spatial domain $\boldsymbol{x}$. The Petersen-Middleton theorem (Petersen and Middleton, 1962) states that the reconstruction of a continuous and band-limited signal from its samples is possible if and only if the spacing of the sampling points is small enough to ensure non-overlapping of the spectrum of the signal with his replicas distributed over the so-called reciprocal lattice (or grid). The latter is defined as the Fourier transform of the specific sampling lattice. The 1D version of this theorem is the well-known Shannon-Nyquist theorem (Shannon, 1984). An application of this theorem to non-uniformly distributed samples, like those measured by remote sensing instruments, is unfeasible due to the lack of periodicity of the sampling points. To circumvent this issue, we adopted the approach suggested by Koch et al. (1983), who defined the random data spacing, Δd, as the equivalent distance that a certain number of samples enclosed in a certain region, $N_{\exp}$, would have if they were uniformly distributed over a structured Cartesian grid. The generalized form of the random data spacing reads:

$$\Delta d(\boldsymbol{r_i}) = \frac{V^{\frac{1}{N}}}{N_{\exp}(\boldsymbol{r_i})^{\frac{1}{N}} - 1}, \tag{7}$$

where V is the volume of the hyper-sphere with radius $R_{\max} = 3\sigma$ centered at the specific grid point and $N_{\exp}$ represents the number of not co-located sample locations included within the hyper-sphere. Then, the Petersen-Middleton theorem for the reconstruction of the generic Fourier mode of half-wavelength $\boldsymbol{\Delta n}$ can be translated as the following constraint:

$$\Delta d(\boldsymbol{r_i}) < \Delta n_p, \ p = 1, 2, ..., N. \tag{8}$$

Violation of the inequality (8) will lead to local aliasing, with the energy content of the under-sampled wavelengths being added to the low-frequency part of the spectrum.

3   LiSBOA assessment through Monte Carlo simulations

The spectral response of the LiSBOA is studied through the Monte Carlo method. The goal of the present section is twofold: validating the analytical response of mean and variance (Eq. (5)) and characterizing the sampling error of the LiSBOA as a function of the random data spacing. For these aims, a synthetic 3D scalar field is generated, while its temporal variability is reproduced locally by randomly sampling a normal probability density function. Specifically, the synthetic scalar field is:

$$f = \left[1 + \sin\left(\frac{\pi}{\Delta n}x\right)\sin\left(\frac{\pi}{\Delta n}y\right)\sin\left(\frac{\pi}{\Delta n}z\right)\right] + \left[1 + \sin\left(\frac{\pi}{\Delta n}x\right)\sin\left(\frac{\pi}{\Delta n}y\right)\sin\left(\frac{\pi}{\Delta n}z\right)\right]^{0.5}\aleph(0,1), \tag{9}$$

where $\aleph$ is a generator of random numbers with normal probability density function with mean value 0 and standard deviation equal to 1. The constant 1 in the two terms on the RHS of Eq. (9) does not affect the LiSBOA response and is introduced to obtain both mean and variance of f equal to the following function:

$$\overline{f} = 1 + \sin\left(\frac{\pi}{\Delta n}x\right)\sin\left(\frac{\pi}{\Delta n}y\right)\sin\left(\frac{\pi}{\Delta n}z\right). \tag{10}$$

265 It is noteworthy that $\overline{f} - 1$ is a monochromatic isotropic function.

An experimental sampling process is mimicked by evaluating the scalar field f through randomly and uniformly distributed samples collected at the locations r_j. The latter are distributed within a cube spanning the range $\pm 10\sigma$ in the three Cartesian directions. The total number of sampling points considered for each realization, N_s, is varied from 500 up to 20,000 to explore the effects of the sample density on the error. The sampling process is repeated L times for each given distribution of N_s points

270 to capture the variability in the field introduced by the operator $\aleph$. The whole procedure can be considered as an idealized LiDAR experiment where a scan including N_s sampling points is performed L times to probe an ergodic turbulent velocity field.

Since the response is only a function of $\Delta n/\sigma$ and m (Eq. (5)), for the spectral characterization of the LiSBOA, the parameter $\Delta n/\sigma$ is varied among the following values: [1, 2, 3, 4, 5]. An implementation of the LiSBOA algorithm for discrete samples

275 is then applied to reconstruct the mean g^m and variance v^m of the scalar field f over a Cartesian structured grid, r_i, with a resolution of 0.25. Figure 1 depicts an example of the reconstruction of the mean scalar field, g^m, and its variance, v^m, from the Monte Carlo synthetic dataset.

For the error quantification, the 95-th percentile of the absolute error calculated at each grid point r_i (AE_{95} hereinafter) is adopted:

280
$$AE_{95}(\Delta n/\sigma, m, N_s, L) = \begin{cases} \text{percentile}_{95} \langle |(g^m - 1) - D^m(\overline{f} - 1)| \rangle_{r_i} & \text{for the mean} \\ \text{percentile}_{95} \langle |(v^m - 1) - D^0(\overline{f} - 1)| \rangle_{r_i} & \text{for the variance.} \end{cases} \tag{11}$$

The AE_{95} quantifies the discrepancy between the outcome of the LiSBOA and the analytical input damped by the theoretical response evaluated over the Cartesian grid. As highlighted in Eq. (11), the expected value of AE_{95} is a function of the half-wavelength over the smoothing parameter, $\Delta n/\sigma$, the number of iterations, m, the number of samples, N_s, and the number of realizations, L. To investigate the link between AE_{95} and the above-mentioned parameters, the Pearson correlation coefficients

285 are analyzed (Table 1). The number of samples N_s, which is inversely proportional to the data spacing Δd (Eq. 7), is the
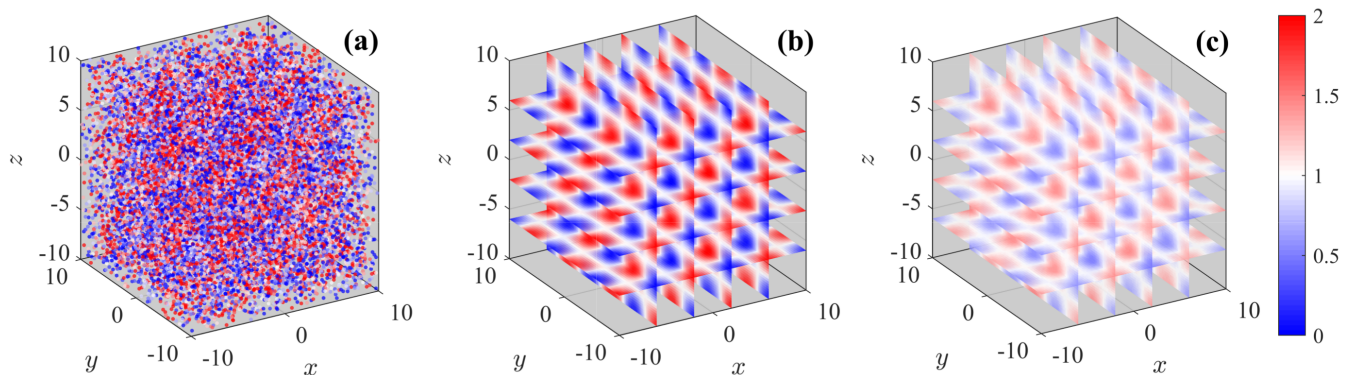
Figure 1. Visualization of the LiSBOA applied to a Monte Carlo simulation of the synthetic field in Eq. (9) for the case with $N_s = 20,000$, $L = 200$, $\Delta n/\sigma = 4$ and $m = 5$: **(a)** samples; **(b)** 3D reconstructed mean field, g^m; **(c)** 3D reconstructed variance, v^m.

variable exhibiting the strongest correlation with the error for both mean and variance. This indicates, as expected, that a larger number of samples for each measurement realization is always beneficial for the estimates of the statistics of the scalar field, f. Furthermore, the negative sign of correlations $\rho(AE_{95}, N_s)$ and $\rho(AE_{95}, \Delta n/\sigma)$, corroborate the hypothesis that the ratio $\Delta d/\Delta n$, i.e. the number of samples per half-wavelength, is the main driving factor for the sampling error (Koch et al., 1983; Barnes, 1994a; Caracena et al., 1984).

The small positive correlation $\rho(AE_{95}, m)$ detected for the mean is due to an amplification of the error occurring during the iterative process (Barnes, 1964). The issue will be discussed more in detail in Sect. 4. For the variance, $\rho(AE_{95}, m)$ is practically negligible, confirming that the response of the higher-order statistics is insensitive to the number of iterations, m. Finally, the negative correlations with L show that the statistical error is inversely proportional to the number of realizations collected. The dependence $\rho(AE_{95}, L)$ is mainly due to the statistical error connected with the temporal sampling and, thus, the number of realizations, L, is progressively increased until convergence of the AE_{95} is achieved. Figure 2 displays the behavior of the error as a function of N_s and L. The values displayed represent the median for all the wavelength and iterations, being the AE_{95} just mildly dependent on these parameters. As Fig. 2 shows, increasing the number of realizations, L, beyond 100 has a negligible effect on the error, thus a final value of $L = 200$ is selected for the remainder of this analysis.
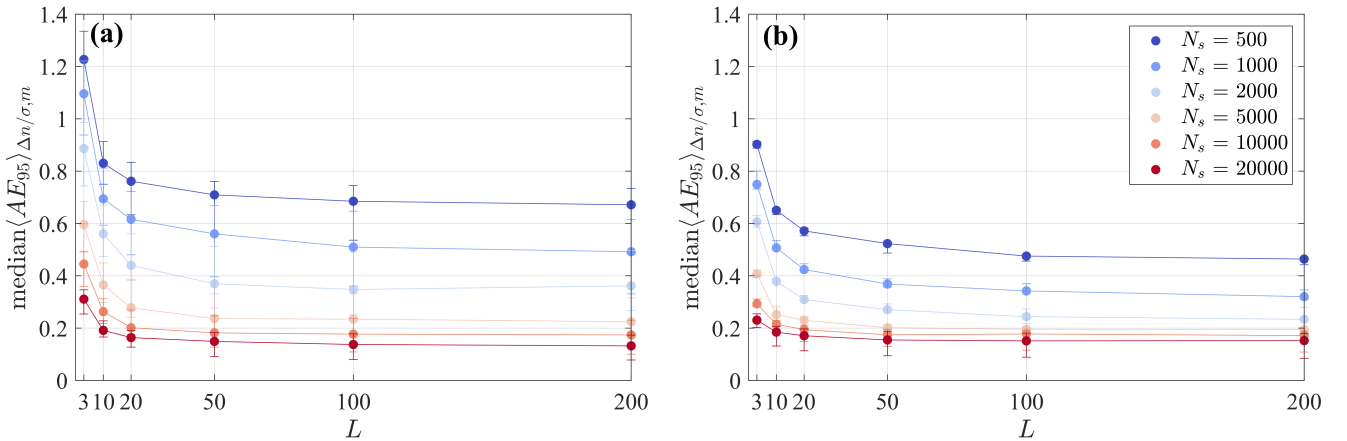
Figure 2. Median of the AE_{95} for all the tested half-wavelengths, $\Delta n/\sigma$, and the number of iterations, m: **(a)** AE_{95} of the mean field, g^m; **(b)** AE_{95} of the variance field, v^m. The error bars span the interquartile range.

Table 1. Pearson correlation coefficient between the AE_{95} of mean and variance and the parameters $\Delta n/\sigma$, m, N_s, L. The values between parenthesis represent the 95% confidence bounds.

	$\Delta n/\sigma$	m	N_s	L
AE_{95} of mean	-0.259 (-0.303, -0.210)	0.257 (0.211, 0.301)	-0.709 (-0.732, -0.684)	-0.171 (-0.217, -0.124)
AE_{95} of variance	-0.069 (-0.117, -0.021)	-0.03 (-0.078, 0.019)	-0.694 (-0.718, -0.668)	-0.206 (-0.251, -0.159)

To verify the analytical response of mean and variance of the scalar field, f, a numerical estimator of the response is defined as the median in space of the ratio between the field reconstructed via LiSBOA and the expected value of the synthetic input, as:

$$\begin{cases} D^m = \text{median}\langle \frac{g^m - 1}{\overline{f} - 1}\rangle_{r_i} & \text{for the mean} \\ D^0 = \text{median}\langle \frac{v^m - 1}{\overline{f} - 1}\rangle_{r_i} & \text{for the variance.} \end{cases} \quad (12)$$

In the calculation of the numerical response through Eq. (12), the influence of the edges is removed by rejecting points closer than $R_{\max}$ to the boundaries of the numerical domain. Furthermore, the zero-crossings of the synthetic sine function ($|\overline{f} - 1| < 0.1$) are excluded to avoid singularities. A comparison between the actual and the theoretical response (Eq. (5)) for several wavelengths of the input function is reported in Fig. 3 for the case with the highest number of samples $N_s = 20,000$. An excellent agreement is observed between the theoretical prediction and the Monte Carlo outcome, which indicates that in the limit of negligible statistical error (large L) and adequate sampling (large N_s and near-uniformly distributed samples) the response approaches the predictions obtained from the developed theoretical framework.

The trend of the response of the mean (Fig. 3a suggests that, for a given wavelength, the same response can be achieved for an infinite number of combinations $\sigma - m$ and, specifically, a larger σ requires a larger number of iterations, m, to achieve a certain response D^m. It is noteworthy that for a smaller number of iterations, m, the slope of the response function is lower. This feature can be beneficial for practical applications for which the LiSBOA response will have small changes for small variations of Δn. However, a lower slope of the response function can be disadvantageous for short-wavelength noise suppression. Figure 3b confirms that the response of the variance, and similarly for higher-order statistics, is not a function of the total number of iterations, m, and is equal to the response of the mean for the 0-th iteration, D^0.
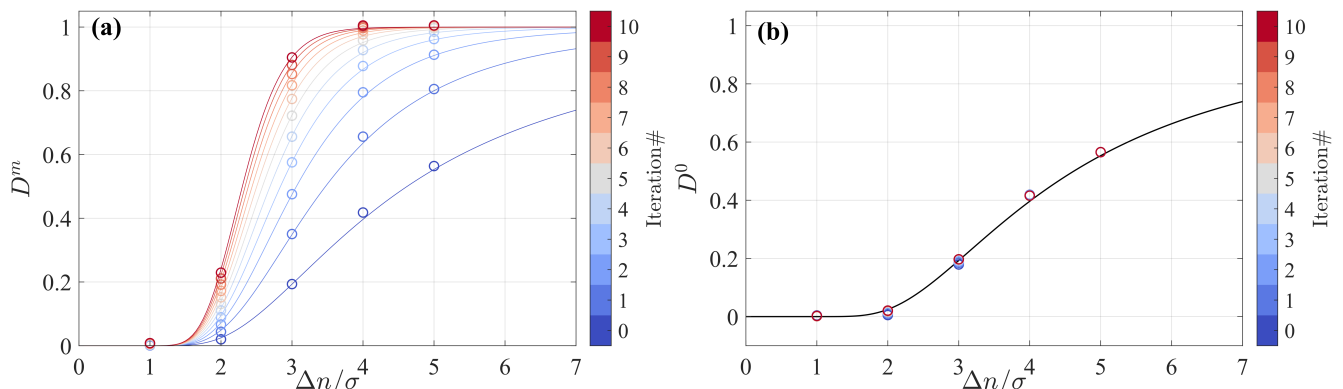
Figure 3. Validation of the 3D theoretical response of the LiSBOA for the case $N_s = 20000$ - $L = 200$: **(a)** mean; **(b)** variance. The circles are the numerical output of the Monte Carlo simulation (Eq. (12)), while the continuous lines represent Eq. (5).

Finally, the link between error and the random data spacing, Δd, is investigated. In Fig. 4, the discrepancy with respect to theory quantified by the AE_{95} is plotted versus the random data spacing normalized by the half-wavelength for a fixed total number of iterations $m = 5$. The values displayed on the x-axis represent the median over all grid points, r_i. This analysis reveals a strong correlation between the normalized random data spacing and the error. This analysis corroborates that, in the limit of negligible statistical error (viz. a high number of realizations, L), uncertainty is mainly driven by the local data density normalized by the wavelength, which is related to the Petersen-Middleton criterion. Indeed, the cases satisfying the Petersen-Middleton constraint (Eq. (8)) are those exhibiting an AE_{95} smaller than $\sim 40\%$ of the amplitude of the harmonic function $\overline{f}$ for both mean and variance. However, if a smaller error is needed, it will be necessary to reduce the maximum threshold value for $\Delta d / \Delta n$.

4 Guidelines for an efficient application of the LiSBOA to wind LiDAR data

An efficient application of the LiSBOA to LiDAR data relies on the appropriate selection of the parameters of the algorithm, namely the fundamental half-wavelengths, $\boldsymbol{\Delta n_0}$, the smoothing parameter, σ, the number of iterations, m, and the spatial discretization of the Cartesian grid, $\boldsymbol{dx}$. Furthermore, the data collection strategy must be designed to ensure adequate sampling of the spatial wavelengths of interest, so that the Petersen-Middleton constraint (Eq. (8)) is satisfied. In this section, we show that the underpinning theory of the LiSBOA, along with an estimate of the properties of the flow under investigation, can guide the optimal design of a LiDAR experiment and evaluation of the statistics for a turbulent ergodic flow. The whole procedure can be divided into three phases: characterization of the flow, design of the experiment, and reconstruction of the statistics from the collected dataset.
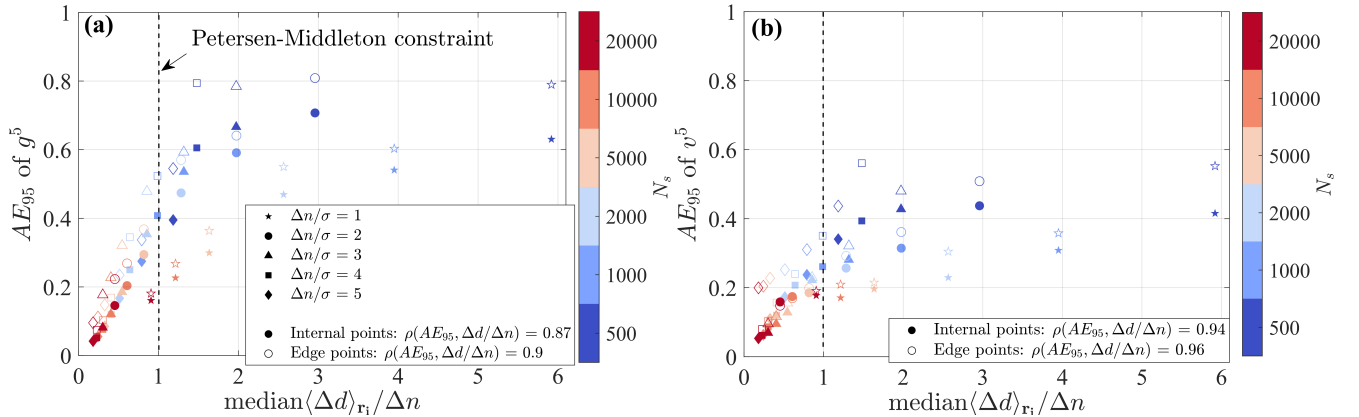
Figure 4. AE_{95} as a function of the random data spacing (Eq. (7)) for the case with $m = 5$, $N = 20000$ and $L = 200$: **(a)** error on the mean; **(b)** error on the variance. The full symbols refer to points not affected by the presence of the finite boundaries of the domain, while the empty symbols are taken within a distance of less than R_{max} from the boundaries.

Firstly, the integral quantities of the flow under investigation required for the application of the LiSBOA need to be estimated, such as extension of the spatial domain of interest, characteristic length-scales, integral time-scale, τ, characteristic temporal variance of the velocity, $\overline{u'^2}$, and expected total sampling time, T, which depends on the typical duration of stationary boundary conditions over the domain. These estimates can be based on previous studies available in the literature, numerical simulations, or preliminary measurements.

Then, it is necessary to define the fundamental half-wavelengths, $\boldsymbol{\Delta n_0}$, which are required for the coordinate scaling (Eq. (6)). It is advisable to impose the fundamental half-wavelengths equal to (or even smaller than) the estimated characteristic length-scales of the smallest spatial features of interest in the flow. This ensures isotropy of the mode associated with the fundamental half-wavelength (and all the modes characterized by the same degree of anisotropy) and guides the selection of the main input parameters of the LiSBOA algorithm, i.e. smoothing parameter, σ, and number of iterations, m. Indeed, $\boldsymbol{\Delta n_0}$ can be considered as the cut-off half-wavelength of the spatial low-pass filter represented by the LiSBOA operator. To this aim, it is necessary to select σ and m to obtain a response of the mean associated with the fundamental mode, $D^m(\boldsymbol{\Delta \tilde{n}_0})$, as close as possible to 1. After the coordinate scaling (Eq. (6)), the response of the fundamental mode is universal and it is reported in Fig. 5. For instance, if we select a response equal to 0.95, then all the points lying on the iso-contour defined by the equality $D^m(\boldsymbol{\Delta \tilde{n}_0}) = 0.95$ give, in theory, the same response for the mean of the scalar field f. This implies that an infinite number of combinations $\sigma - m$ allow obtaining a response of the mean equal to the selected value. However, with increasing σ the response at the 0-th iteration, $D^0(\boldsymbol{\Delta \tilde{n}_0})$, reduces, which indicates a lower response for higher-order statistics. For the LiSBOA application, the following aspects should be also considered:

- the smaller σ, the smaller the radius of influence of the LiSBOA, $R_{\max}$, and, thus, the lower the number of samples averaged per grid node, $N_{\exp}$, and the greater the statistical uncertainty;

- an excessively large m can lead to overfitting of the experimental data and noise amplification (Barnes, 1964);

- the higher m, the higher the slope of the response function (see Fig. 3), which improves the damping of high-frequency noise, but it produces a larger variation of the response of the mean with different spatial wavelengths;

- the radius of influence $R_{\max}$ (and therefore σ) can affect the data spacing Δd in case of non-uniform data distribution.

Few handy combinations of smoothing parameter and total iterations for $D^m(\boldsymbol{\Delta \tilde{n}_0}) = 0.95$ are provided in Table 2. As mentioned above, all these $\sigma - m$ pairs allow achieving roughly the same response for the mean, while the response for the higher-order statistics reduces with an increasing number of iterations, m.

As a final remark about the selection of $\boldsymbol{\Delta n_0}$, we should consider that, if the fundamental half-wavelength is too large compared to the dominant modes in the flow, small-scale spatial oscillations of f will be smoothed out during the calculation of the mean, with consequent underestimated gradients and incorrect estimates of the high-order statistics due to the dispersive stresses (Arenas et al., 2019). On the other hand, the selection of an overly small $\boldsymbol{\Delta n_0}$, would require an excessively fine data spacing to satisfy the Petersen-Middleton constraint (Eq. (8)), which may lead to an overly long sampling time or even exceed the sampling capabilities of the LiDAR.
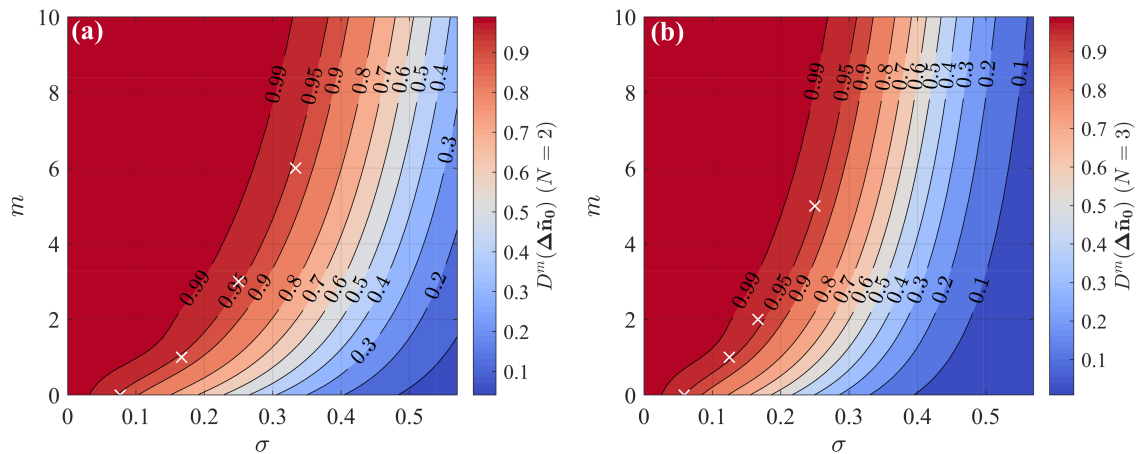
Figure 5. Response of the fundamental mode in the scaled coordinates as a function of the number of iterations and the smoothing parameter: **(a)** two-dimensional LiSBOA; **(b)** three-dimensional LiSBOA. The white crosses indicate the pairs $\sigma - m$ provided in Table 2.

Table 2. Selected combinations of σ and m to achieve a $\sim$95% recovery of the mean of the selected fundamental half-wavelength and associated response of the higher-order moments (HOM).

	$N = 2$				$N = 3$		
σ	m	D^m (mean)	D^0 HOM	σ	m	D^m (mean)	D^0 (HOM)
1/3	6	0.942	0.334	1/4	5	0.952	0.397
1/4	3	0.955	0.540	1/6	2	0.961	0.663
1/6	1	0.942	0.76	1/8	1	0.957	0.793
1/13	0	0.943	0.943	1/17	0	0.950	0.950

The optimal LiDAR scanning strategy aimed to characterize atmospheric turbulent flows implies finding a trade-off between a sufficiently fine data spacing, which is quantified through Δd in the present work (Eq.(7)), and an adequate number of time realizations, L, to reduce the temporal statistical uncertainty. Considering a total sampling period T, for which statistical stationarity can be assumed, and a pulsed LiDAR that scans N_r points evenly spaced along the LiDAR laser beam, with a range gate Δr and accumulation time τ_a, the total number of collected velocity samples is then equal to $N_s = N_r \cdot T/\tau_a$. The angular resolution of the LiDAR scanning head in azimuth ($\Delta\theta$, for PPIs) or elevation ($\Delta\beta$, for RHIs) or both axis (for volumetric scans), can be selected to modify the angular spacing between consecutive line-of-sights (i.e. the data spacing) and the total sampling period for a single scan, τ_s (i.e. the number of realizations, L).

The design of a LiDAR scan aiming to reconstruct turbulent statistics of an ergodic flow through LiSBOA can be formalized as a two-objective (or Pareto front) optimization problem. The first cost function of the Pareto front, which is referred to as ϵ^I, is the percentage of grid nodes for which the Petersen-Middleton constraint applied to the smallest half-wavelength of interest

(i.e. Δn_0), is not satisfied. In respect to the scaled reference frame, this can be expressed as:

$$\epsilon^I(\Delta\theta, \Delta\beta, \sigma) = \frac{\sum_{i=1}^{N_i} [\Delta\tilde{d} > 1]}{N_i}, \tag{13}$$

where the square brackets are Iverson brackets and N_i is the total number of nodes in the Cartesian grid, $\boldsymbol{r_i}$. For a more conservative formulation, it is recommended to reject all the points with a distance smaller than $R_{\max}$ from an under-sampled grid node, i.e. with $\Delta\tilde{d} > 1$. This condition will ensure that the statistics are based solely on regions that are adequately sampled. The cost function ϵ^I depends not only on the angular resolution but also on $R_{\max}$, which is equal to 3σ in this work. In general, increasing σ results in a larger number of samples considered for the calculation of the statistics at each grid point $\boldsymbol{r_i}$ and, thus, in a reduction of ϵ^I. Therefore, a larger σ entails a larger percentage of the spatial domain fulfilling the Petersen-Middleton constraint. The smoothing parameter, σ, also plays a fundamental role in the response of higher-order statistical moments. Specifically, if the reconstruction of the variance or higher-order statistics is important, the response $D^0(\Delta\tilde{n}_0)$ should be included in the Pareto front analysis as an additional constraint.

The second cost function for the optimal design of LiDAR scans, ϵ^{II}, is equal to the standard deviation of the sample mean, which, for an autocorrelated signal, is (Bayley and Hammersley, 1946):

$$\epsilon^{II}(\Delta\theta, \Delta\beta) = \sqrt{\overline{u'^2}}\sqrt{\frac{1}{L} + \frac{2}{L^2}\sum_{p=1}^{L-1}(L-p)\,\rho_p} \sim \sqrt{\overline{u'^2}}\sqrt{\frac{1}{L} + \frac{2}{L^2}\sum_{p=1}^{L-1}(L-p)\,e^{-\frac{\tau_s}{\tau}p}}, \tag{14}$$

where ρ_p is the autocorrelation function at lag p, τ is the integral time-scale and the approximation is based on George et al. (1978). The velocity variance, $\overline{u'^2}$, and the autocorrelation, ρ_p, are functions of space; however, to a good degree of approximation, they can be replaced by a representative value and considered as uniform in space. Figure 6 shows the standard deviation of the sample mean normalized by the standard deviation of the velocity as a function of the number of realizations, L, and for different integral time-scales, τ. It is noteworthy that the standard deviation of the sample mean represents the uncertainty of the time-average of each measurement point, $\boldsymbol{r_j}$, while the final uncertainty of the mean field at the grid nodes $\boldsymbol{r_i}$ is generally reduced due to the spatial averaging process intrinsic to the LiSBOA. It is noteworthy that the estimates of the statistical error obtained through the LiSBOA do not consider other sources of error, such as accuracy of the instruments and spatial averaging due to the LiDAR measuring process (Rye and Hardesty, 1993; O'Connor et al., 2010; Puccioni and Iungo, 2020). Eventually, other error estimates can be coupled with the sampling error estimated through the LiSBOA for a more comprehensive error analysis (Wheeler and Ganji, 2010b).

The whole procedure for the design of a LiDAR scan and retrieval of the statistics is reported in the flow chart of Fig. 7. Summarizing, from a preliminary analysis of the velocity field under investigation, we estimate the maximum total sampling time, T, the characteristic integral time-scale, τ, the characteristic velocity variance, $\overline{u'^2}$, the fundamental half-wavelengths Δn_0. This information, together with the settings of the LiDAR (namely, the accumulation time, τ_a, the number of points per beam, N_r, and the gate length, Δr), allow for generating the Pareto front as a function of $\Delta\theta$ and/or $\Delta\beta$, and for different values of σ. Based on the specific goals of the LiDAR campaign in terms of coverage of the selected domain (i.e. ϵ^I), statistical significance of the data (i.e. ϵ^{II}) and, eventually, response of the higher-order statistical moments (i.e. $D^0(\Delta\tilde{n}_0)$), the LiSBOA
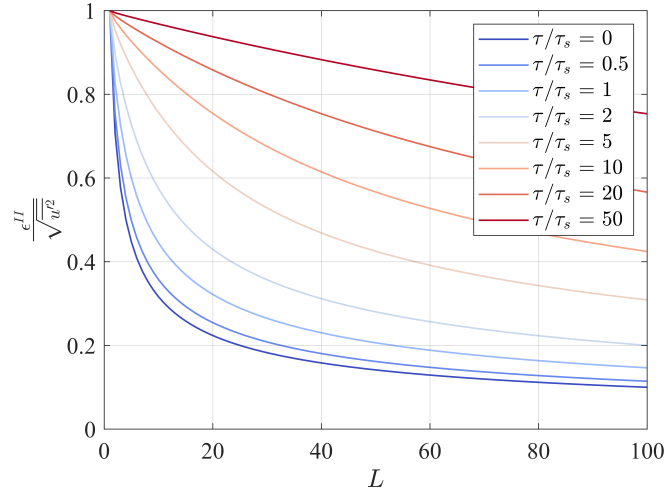
Figure 6. Standard deviation of the sample mean normalized by the standard deviation of velocity as a function of the number of realizations, L, and for different values of the ratio between the integral time-scale and the sampling time, τ/τ_s.
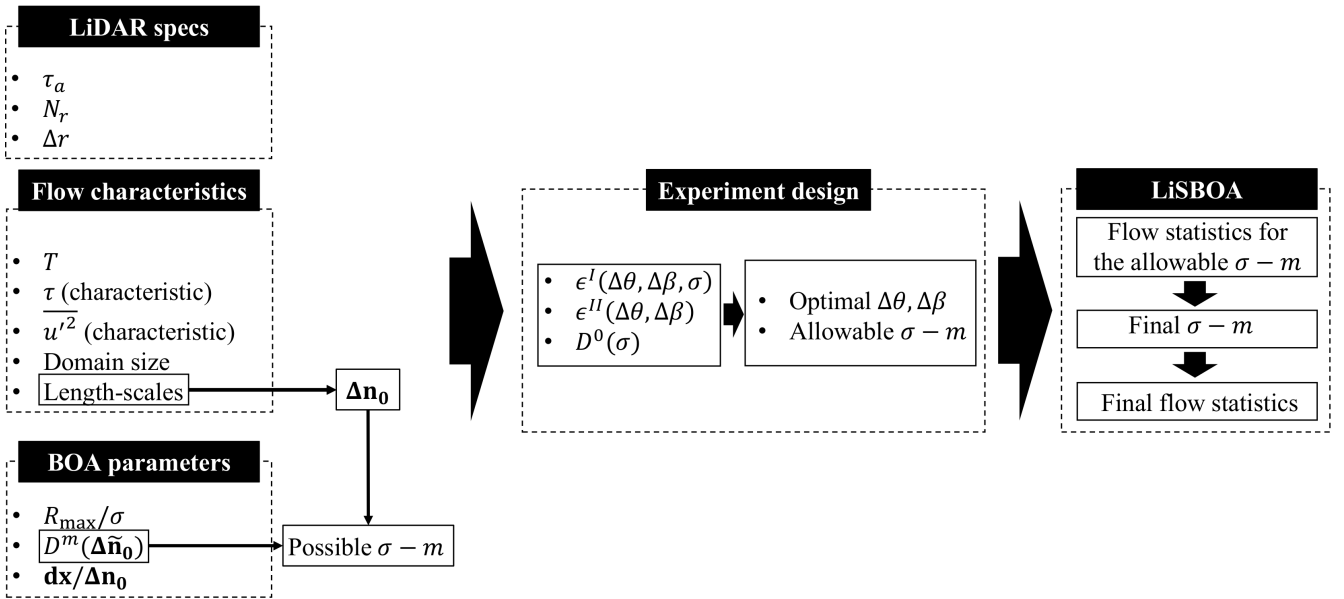
Figure 7. Schematic of the LiSBOA procedure for the optimal design of LiDAR scans and reconstruction of the statistics for a turbulent ergodic flow.

user should select the optimal angular resolution, $\Delta\theta$ and/or $\Delta\beta$, and the set of allowable σ values. Due to the above-mentioned non-ideal effects on the LiSBOA, the selection of σ is finalized during the post-processing phase when the LiDAR dataset is available and the statistics can be calculated for different pairs of $\sigma - m$ values. For the resolution of the Cartesian grid, Koch

et al. (1983) suggested that it should be chosen as a fraction of the data spacing, which, in turn, is linked to the fundamental half-wavelength. The same author suggested a grid spacing included in the range $dx \in [\Delta n_0/3, \Delta n_0/2]$. In this work, we have used $dx = \Delta n_0/4$, which ensures a good grid resolution with acceptable computational costs.

By following the steps outlined in the present section, the mean, variance, or even higher-order statistical moments of the velocity field can be accurately reconstructed for the wavelengths of interest. It is worth mentioning that the LiSBOA of wind LiDAR data should always be combined with a robust quality-control process of the raw measurements. Indeed, the space-time averaging operated by the LiSBOA makes the data analysis sensitive to the presence of data outliers, which need to be identified and rejected beforehand to prevent contamination of the final statistics. The interested reader is referred to Manninen et al. (2016); Beck and Kühn (2017); Vakkari et al. (2019) for more information on quality control of LiDAR data. On a final note, for applications of the LiSBOA, the uncontrollable environmental conditions and the uncertainty in the flow characteristics needed as input of the LiSBOA may pose some challenges, which will be discussed more in detail in Sect. 6.

5   LiSBOA validation against virtual LiDAR data

The LiSBOA algorithm is applied to a synthetic dataset generated through the virtual LiDAR technique to assess accuracy in the calculation of statistics for a wind turbine wake probed through a scanning LiDAR installed at the turbine nacelle. For this purpose, a simulator of a scanning Doppler pulsed wind LiDAR is implemented to extract the line-of-sight velocity from a numerical velocity field produced through high-fidelity large-eddy simulations (LES). Due to their simplicity and low computational costs, LiDAR simulators have been widely used for the assessment of post-processing algorithms of LiDAR data and scan design procedures (Mann et al., 2010; Stawiarski et al., 2015; Lundquist et al., 2015; Mirocha et al., 2015).

As a case study, we use the LES dataset of the flow past a single turbine with the same characteristics of the 5-MW NREL reference wind turbine (Jonkman et al., 2009). The rotor is three-bladed and has a diameter $D = 126$ m. The tip-speed ratio of the turbine is set to its optimal value of 7.5. A uniform incoming wind with freestream velocity of $U_\infty = 10$ m s^{-1} and turbulence intensity of 3.6% is considered. The rotor is simulated through an actuator disk with rotation, while the drag of the nacelle is taken into account using an immersed boundary method (Ciri et al., 2017). More details on the LES solver can be found in Santoni et al. (2015). The computational domain has dimensions ($L_x \times L_y \times L_z = 12D \times 6D \times 6D$) in the streamwise, spanwise, and vertical directions, respectively, and it is discretized with $960 \times 256 \times 300$ uniformly spaced grid points, respectively, resulting in a spacing of $dx = 0.0125D$, $dy = 0.025D$ and $dz = 0.0202D$. A radiative condition is imposed at the outlet (Orlanski, 1976), while periodicity is applied in the spanwise direction. For the sake of generality, a uniform incoming wind is generated by imposing freeslip conditions at the top and bottom of the numerical domain. Ergodic velocity vector fields are available for a total time of $T = 750$ s.

For the estimation of the flow characteristics necessary for the scan design, the azimuthally-averaged mean and standard deviation of streamwise velocity, as well as the integral time-scale are considered (Fig. 8). The use of cylindrical coordinates is justified by the axisymmetry of the statistics of the wake velocity field generated by a turbine operating in a uniform velocity field (Iungo et al., 2013a; Viola et al., 2014; Ashton et al., 2016).
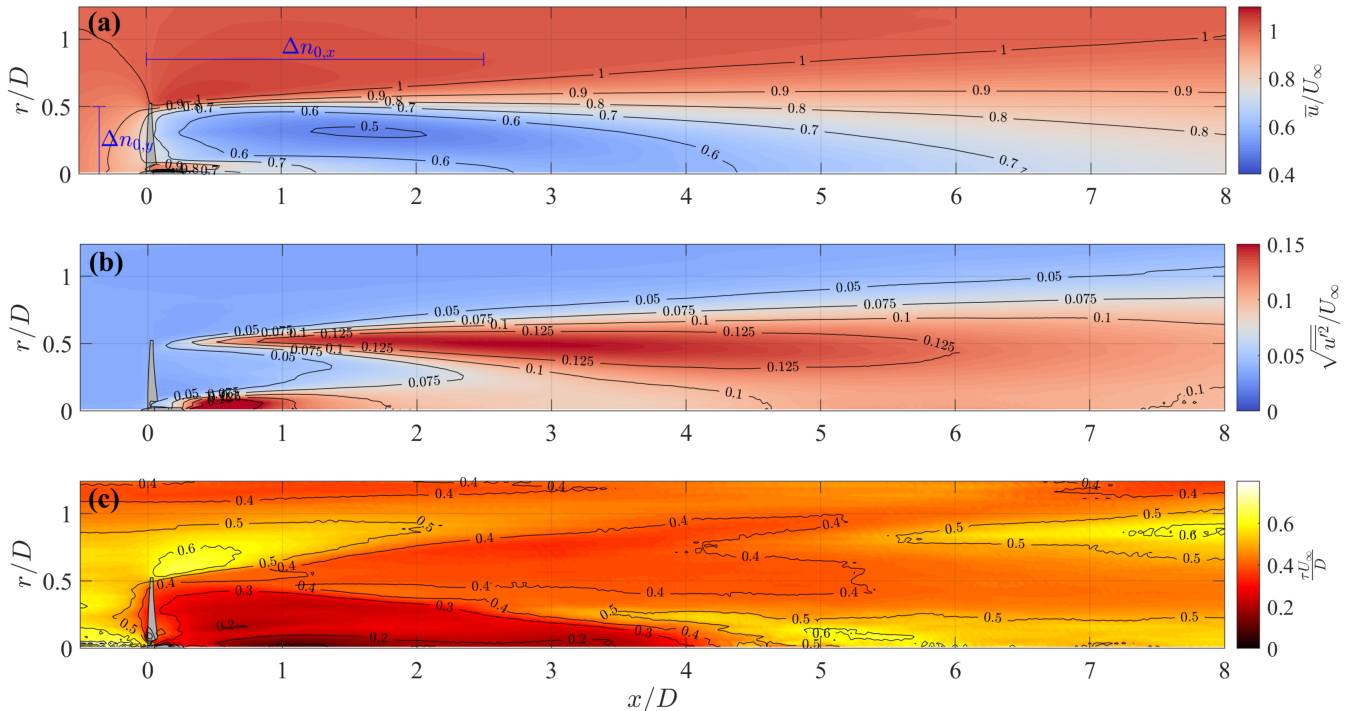
Figure 8. Azimuthally-averaged statistics of the LES streamwise velocity field: **(a)** mean value; **(b)** standard deviation; **(c)** integral time scale.

The streamwise LES velocity field shows the presence of a higher-velocity jet surrounding the nacelle, while $\overline{u}/U_\infty$ exhibits a clear minimum placed at $y/D \sim 0.25$ (Fig. 8a). These flow features are consistent with the double-Gaussian velocity profile typically observed in the near-wake region (Aitken and Lundquist, 2014). In Fig. 8b, the standard deviation of the streamwise 450 velocity has high values in the very near-wake ($x/D < 1$) in the proximity of the rotor axis, which is most probably connected with the vorticity structures generated in proximity of the rotor hub and their dynamics (Iungo et al., 2013a; Viola et al., 2014; Ashton et al., 2016). Similarly, enhanced values of the velocity standard deviation occur at the wake boundary ($r/D \approx 0.5$), which are connected with the formation and dynamics of the helicoidal tip vortices (Ivanell et al., 2010; Debnath et al., 2017c). A peak of $\sqrt{\overline{u'^2}}/U_\infty$ is observed around ($x/D \approx 3$), which can be considered as the formation length of the tip vortices. The 455 integral time-scale is evaluated integrating the sample biased autocorrelation function of the time series of u up to the first zero-crossing (Zieba and Ramza, 2011). The integral time-scale is generally smaller within the wake than for the typical values observed in the freestream, which is consistent with the smaller dimensions of the wake vorticity structures compared to the larger energy-containing structures present in the incoming turbulent wind.

To reconstruct the mentioned flow features, the fundamental half-wavelengths in the spanwise and vertical directions selected 460 for this application of the LiSBOA are $\Delta n_{0,y} = \Delta n_{0,z} = 0.5D$, which allows retrieving modes as small as the rotor blade in the cross-stream direction, which are typically observed in the near wake (Aitken and Lundquist, 2014; Santoni et al., 2017).

Furthermore, considering the streamwise elongation of the isocontours of the flow statistics shown in Fig. 8a, a conservative value of the fundamental half-wavelength in the x-direction $\Delta n_{0,x} = 2.5D$ is selected. This information could have been inferred also from previous field measurements (e.g. Chamorro and Porté-Agel (2010); Abkar and Porté-Agel (2013); Zhan et al. (2019)).

The availability of the numerical flow allows testing the relevance of the selected $\boldsymbol{\Delta n_0}$ by evaluating the 3D energy spectrum of $\overline{u}/U_\infty$ and $\overline{u'^2}/U_\infty^2$ in the physical and in the scaled reference frame (Eq. (6)). The spectra are azimuthally averaged by exploiting the axisymmetry of the wake. The spectra in the physical reference frame (Figs. 9a and b) reveal the clear signature of a streamwise elongation of the energy-containing scales for both velocity mean and variance, with the energy being spread over a larger range of frequencies in the radial direction compared to the streamwise direction. After the scaling (Figs. 9c and d), the spectra become more isotropic in the spectral domain, namely the energy is distributed equally along the $\tilde{k}_x$ and $\tilde{k}_r$ axes. In Fig. 9c, the blue dashed line represents the intersection with the $\tilde{k}_x$ - $\tilde{k}_r$ plane of the spherical isosurface that in the
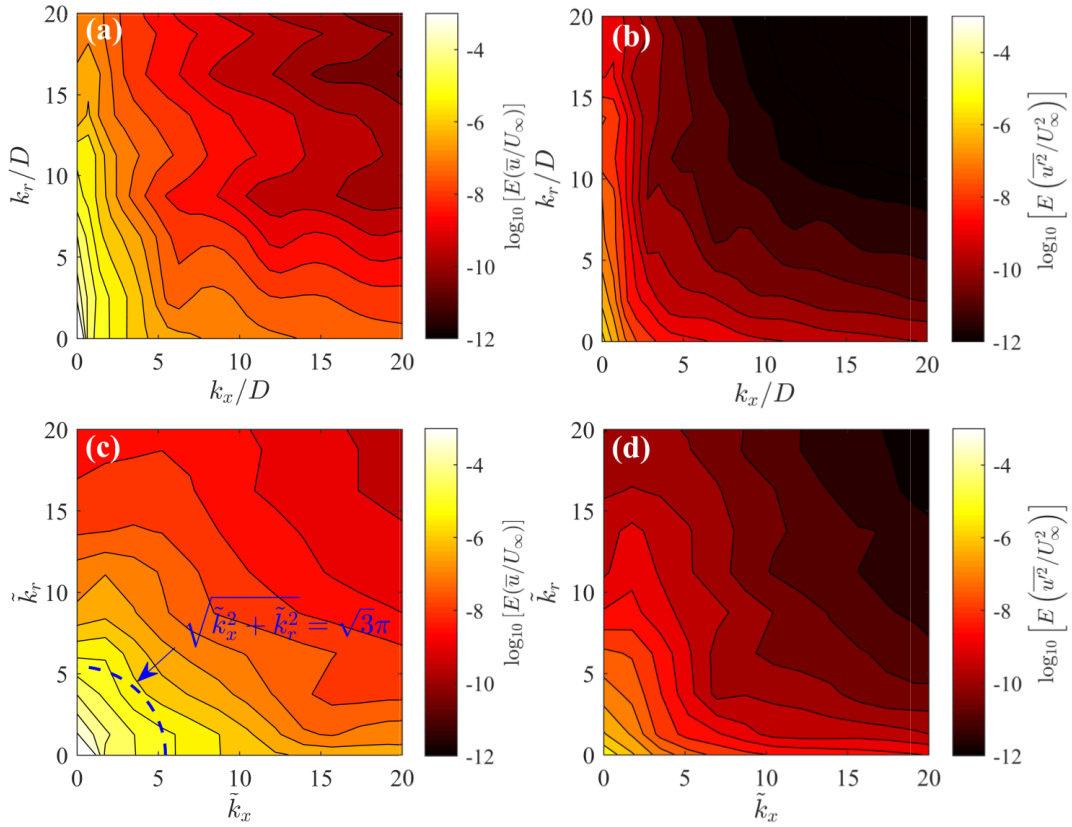
Figure 9. Azimuthally-averaged energy spectra of the LES velocity fields : **(a)** mean streamwise velocity on the physical domain; **(b)** variance of streamwise velocity on the physical domain; **(c)** mean streamwise velocity on the scaled domain; **(d)** variance of streamwise velocity on the scaled domain. The blue dashed line indicates wavenumbers reconstructed with response equal to $D^m(\boldsymbol{\Delta \tilde{n}_0}) = 0.95$.

wavenumber space is characterize by $D^m(\Delta\tilde{\boldsymbol{n}}_0) = 0.95$. All the modes contained within that sphere are reconstructed with a response $D^m > 0.95$, while higher-frequency features lying outside will be damped. Numerical integration of the 3D energy spectrum shows that 94% of the total spatial variance of the mean is contained within that sphere, which ensures that the energy-containing modes in the mean flow are adequately reconstructed with the selected parameters.

The analysis of the flow statistics reported in Fig. 8 enables estimates of flow parameters needed as input for the LiSBOA. For instance, the wake region is characterized by $\sqrt{\langle\overline{u'^2}\rangle}/U_\infty \approx 0.1$ and $\tau U_\infty/D \approx 0.4$ ($\tau \approx 5$ s).

A main limitation of LiDARs is represented by the spatio-temporal averaging of the velocity field, which is connected with the acquisition process. Three different types of smoothing mechanisms can occur during the LiDAR sampling: the first is the averaging along the laser beam direction within each range gate, which has been commonly modeled through the convolution of the actual velocity field with a weighting function within the measurement volume (Smalikho, 1995; Frehlich, 1997; Sathe et al., 2011). The second process is the time-averaging associated with the sampling period required to achieve a back-scattered signal with adequate intensity (O'Connor et al., 2010; Sathe et al., 2011), while the last one is the transverse averaging (azimuth-wise or elevation-wise) occurring in case of a scanning LiDAR operating in continuous mode (Stawiarski et al., 2013). These filtering processes lead to significant underestimation of the turbulence intensity (Sathe et al., 2011), overestimation of integral length-scales (Stawiarski et al., 2015), and damping of energy spectra for increasing wavenumbers (Risan et al., 2018; Puccioni and Iungo, 2020).

Three versions of a LiDAR simulator are implemented for this work: the simplest one is referred to as ideal LiDAR, which samples the LES velocity field at the experimental points through a nearest-neighbor interpolation. This method minimizes the turbulence damping while retaining the geometry of the scan and the projection of the wind velocity vector onto the laser beam direction. The second version of the LiDAR simulator reproduces a step-stare LiDAR, namely the LiDAR scans for the entire duration of the accumulation time at a fixed direction of the LiDAR laser beam. Two filtering processes take place for this configuration: beam-wise convolution and time averaging. To model the beam-wise average, the retrieval process of the Doppler LiDAR is reproduced using a spatial convolution (Mann et al., 2010):

$$u_{\text{LOS}}(\boldsymbol{x},t) = \int\limits_{-\infty}^{\infty} \phi(s)\boldsymbol{n}\cdot\boldsymbol{u}(\boldsymbol{x}+\boldsymbol{n}s,t)ds, \tag{15}$$

where $\boldsymbol{n}$ is the LiDAR laser-beam direction, $\boldsymbol{u}$ is the instantaneous velocity vector and the dot indicates scalar product. A triangular weighting function $\phi(s)$ was proposed by Mann et al. (2010):

$$\phi(s) = \begin{cases} \frac{\Delta r/2-|s|}{\Delta r^2/4} & \text{if } |s| < \Delta r/2 \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

where Δr is the gate length. The former expression is valid assuming matching time windowing, i.e. gate length equal to the pulse width, and the velocity value is retrieved based on the first momentum of the back-scattering spectrum. Despite its simplicity, Eq. (16) has shown to estimate realistic turbulence attenuation due to the beam-wise averaging process of a pulsed Doppler wind LiDAR (Mann et al., 2009). Furthermore, time-averaging occurs due to the accumulation time necessary for the

LiDAR to acquire a velocity signal with sufficient intensity and, thus, signal-to-noise-ratio. This process is modeled through a window average within the acquisition interval of each beam. For the sampling of the LES velocity field in space and time, a nearest-neighbor interpolation method is used.

The third version of the LiDAR simulator mimics a pulsed LiDAR operating in continuous mode and performing PPI scans, where, in addition to the beam-wise convolution and time-averaging, azimuth-wise averaging occurs due to the variation of the LiDAR azimuth angle of the scanning head during the scan. The latter is taken into account by adding to the time average an azimuthal averaging among all data points included within the following angular sector:

$$
\begin{cases}
|\theta - \theta_p| < \Delta\theta/2 \\
|\beta - \beta_p| < \sin^{-1}\left(\frac{\Delta z}{2 r_p}\right)
\end{cases}
\tag{17}
$$

where r is the radial distance from the emitter, while θ and β are the associated azimuth and elevation angles, respectively. The subscript p refers to the p-th LiDAR data point. Following the suggestions by Stawiarski et al. (2013), the out-of-plane thickness, Δz, is considered equal to the length of the diagonal of a cell of the computational grid.

It is noteworthy that the accuracy estimated through the present analysis only includes error due to the sampling in time and space, and data retrieval. Other error sources, such as accuracy of the instrument (Rye and Hardesty, 1993; O'Connor et al., 2010), are not included and should be coupled to the LiSBOA estimates for a more general error quantification (Wheeler and Ganji, 2010b).

Figure 10a shows a snapshot of the streamwise velocity field over the horizontal plane at hub height obtained from the LES. The respective data of the radial velocity obtained from the three versions of the LiDAR simulator by considering a scanning pulsed wind LiDAR deployed at the turbine location and at hub height highlight the increased spatial smoothing of the radial velocity field by adding the various averaging processes connected with the LiDAR measuring process, namely beam-wise, temporal and azimuthal averaging (Fig. 10).

The application of the LiSBOA requires to provide technical specifications of the LiDAR, specifically accumulation time, τ_a, number of gates, N_r, and gate length, Δr. For this work, these parameters are selected based on the typical settings of LiDARs Windcube 200S and StreamLine XR (El-Asha et al., 2017; Zhan et al., 2019, 2020), namely $\tau_a = 0.5$ s, $N_r = 39$ and $\Delta r = 25$ m. Furthermore, to probe the wake region, a volumetric scan including several PPI scans with azimuth and elevation angles uniformly spanning the range $\pm 10°$ with a constant angular resolution, $\Delta\theta = \Delta\beta$, is selected, while the virtual LiDAR is placed at the turbine hub.

With the information provided about the flow under investigation and the LiDAR system, it is possible to draw the Pareto front for the optimization of the LiDAR scan as a function of the angular resolution of the LiDAR scanning head, $\Delta\theta$, and the smoothing parameter of the LiSBOA, σ, as shown in Fig. 11 for the case under investigation.

For the optimization of the LiDAR scan, the LiDAR angular resolution, $\Delta\theta$, is evenly varied, for a total number of 7 cases, from $0.75°$, corresponding to a total number of scans $L = 2$, to $4°$, which corresponds to $L = 41$. The four values of σ recommended in Table 2 to achieve a response of the mean $D^m(\Delta\tilde{n}_0) = 0.95$ are considered here. In Fig. 11, markers indicate the different σ and, thus, the response of high-order statistical moments, $D^0(\Delta\tilde{n}_0)$. The Pareto front shows that increasing $\Delta\theta$
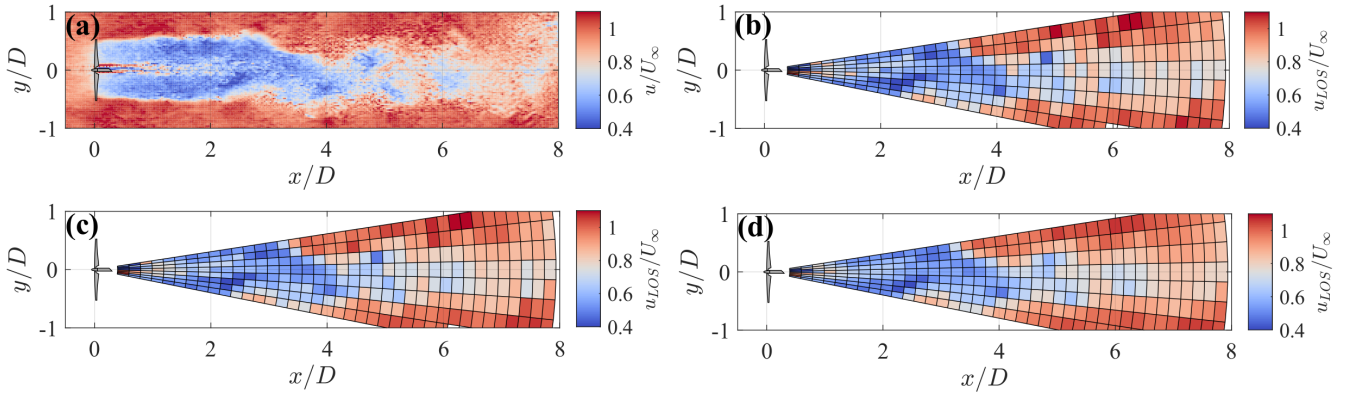
Figure 10. Snapshot at the hub-height horizontal plane of the wake generated by the 5-MW NREL reference wind turbine: **(a)** LES stream-wise velocity; **(b)** ideal virtual LiDAR with angular resolution $\Delta\theta = 2.5°$, zero elevation, accumulation time $\tau_a = 0.5$ s, gate length $\Delta r = 25$ m; **(c)** step-stare virtual LiDAR (same settings); **(d)** continuous mode virtual LiDAR (same settings).
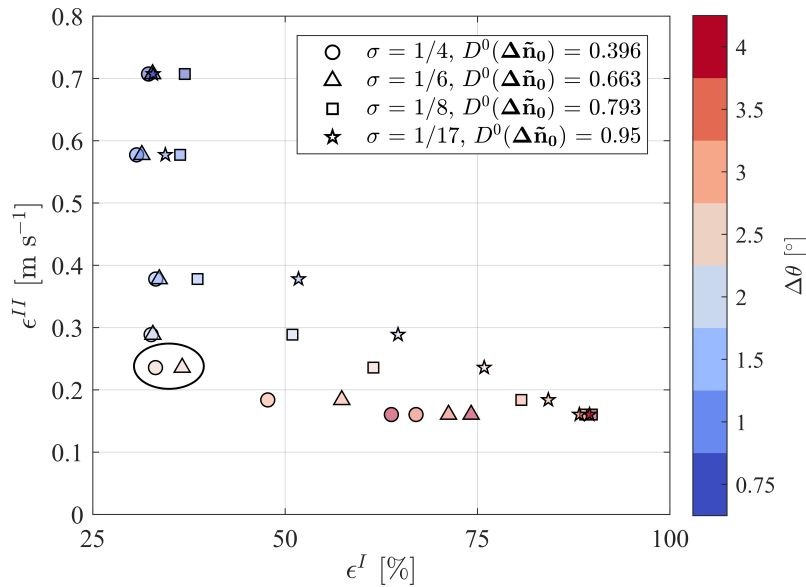
Figure 11. Pareto front for the design of the optimal LiDAR scan for the LES dataset. The circle indicates the selected optimal configurations.

from 0.75° up to 2.5° drastically reduces the uncertainty on the mean (ϵ^{II}) by roughly 70 %, but does not affect significantly data loss consequent to the enforcement of the Petersen-Middleton constraint (ϵ^{I}). For larger angular resolutions, the statistical significance improves just marginally, but at the cost of a relevant data loss. For $\Delta\theta \geq 2°$, in particular, ϵ^{I} becomes extremely sensitive to σ, with the most severe data loss occurring for small σ (i.e. small $R_{\max}$). The Pareto front also shows that to achieve a higher response for the higher-order statistics, $D^0(\mathbf{\Delta\tilde{n}_0})$, generally entails an increased data loss and/or statistical

uncertainty of the mean. This analysis suggests that the optimal LiDAR scan for the reconstruction of the mean velocity field should be performed with $\Delta\theta = 2.5°$ and $\sigma = 1/4$ or $1/6$.

Virtual LiDAR simulations are performed for all the values of angular resolution utilized in the Pareto front reported in Fig. 11. The streamwise component is estimated from the line-of-sight velocity through an equivalent velocity approach (Zhan et al., 2019). The latter states that, for small elevation angles (i.e. $\beta \ll 1$) and under the assumptions of negligible vertical velocity compared to the horizontal component (i.e. $|w| \ll \sqrt{u^2+v^2}$) and uniform wind direction, θ_w, a proxy for the streamwise velocity can be calculated as:

$$u \sim \frac{u_{\text{LOS}}}{cos(\theta - \theta_w)cos\beta}. \tag{18}$$

The mean velocity and turbulence intensity are reconstructed through the LiSBOA. The maximum error is quantified through the 95-th percentile of the absolute error, AE_{95}, using as reference the LES statistics interpolated on the LiSBOA grid.

Fig. 12 reports the AE_{95} for the flow statistics for all the virtual experiments. The error for the mean field (Figs. 12a-c) is mostly governed by the angular resolution, with a higher error occurring for slower scans. This is a clear consequence of the increased statistical uncertainty due to the limited number of scan repetitions, L, that are achievable for small $\Delta\theta$ values and a
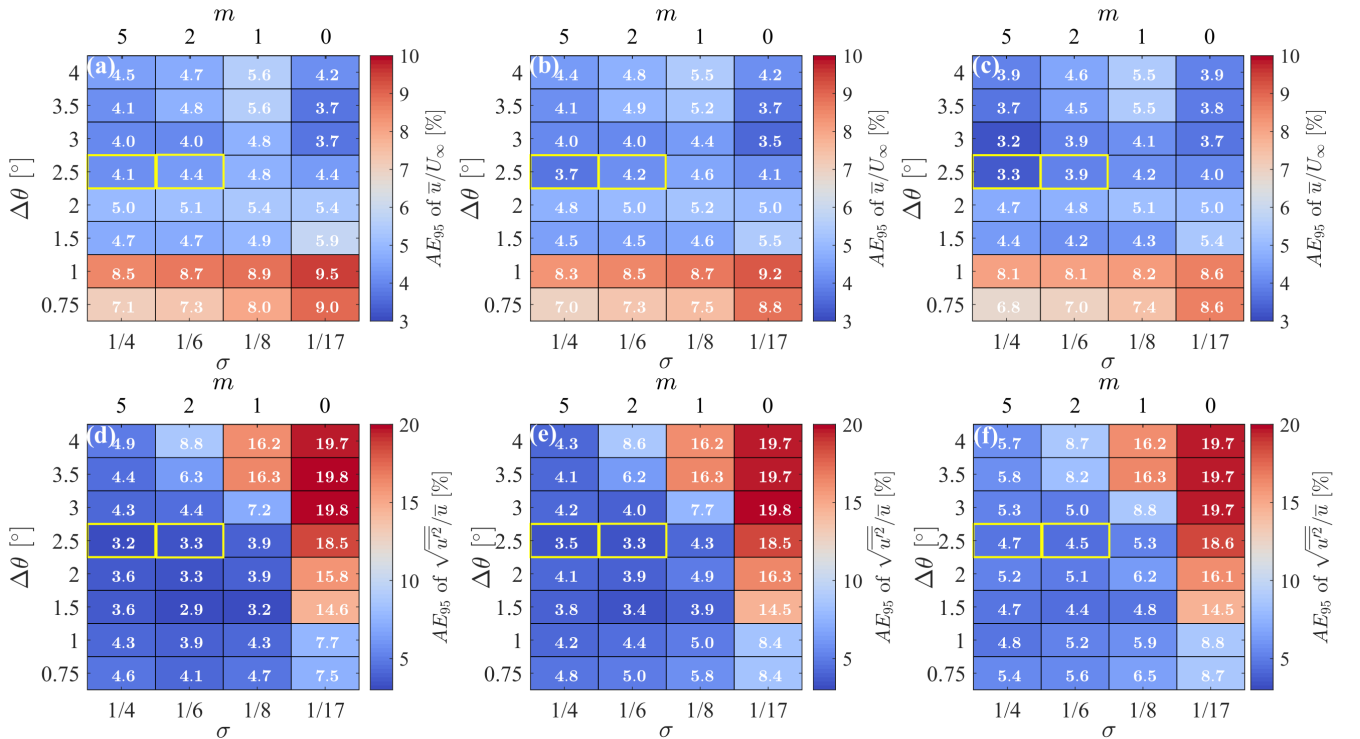
Figure 12. Error analysis of the LiSBOA applied to virtual radial velocity fields: **(a)** and **(d)** ideal LiDAR; **(b)** and **(e)** step-stare LiDAR; **(c)** and **(f)** continuous LiDAR; **(a)**, **(b)** and **(c)** mean streamwise velocity; **(d)**, **(e)** and **(f)** streamwise turbulence intensity. The optimal configurations are highlighted in yellow.

fixed total sampling period, T, while AE_{95} stabilizes for $\Delta\theta \geq 2.5°$. The trend of the AE_{95} for $\overline{u}/U_\infty$ with the pair smoothing parameter-number of iterations, $\sigma - m$, is less significant since the theoretical response of the fundamental mode is ideally equal for all the four cases. Conversely, the error on the turbulence intensity (Figs. 12d-f) shows low sensitivity to the angular resolution but a steep increase for small σ values, which is due to the reduction of the radius of influence, $R_{\max}$, and the number of points averaged per grid node.

From a more technical standpoint, the error on the mean velocity field, $\overline{u}/U_\infty$, appears to be relatively insensitive to the type of LiDAR scan, with the spatial and temporal filtering operated by the step-stare and continuous LiDAR being even beneficial in some cases. In contrast, the error on the turbulence intensity exhibits a more consistent and opposite trend, with the continuous LiDAR showing the most severe turbulence damping. This feature has been extensively documented in previous studies, see e.g. Sathe et al. (2011).

This error analysis confirms that the optimal configurations selected through the Pareto front (i.e. $\Delta\theta = 2.5°$, $\sigma = 1/4 - m = 5$ and $\sigma = 1/6 - m = 2$) are arguably optimal in terms of accuracy (AE_{95} of $\overline{u}/U_\infty = 3.3 - 4.1\%$ and $3.9 - 4.4\%$, AE_{95} of $\sqrt{\overline{u'^2}}/\overline{u} = 3.2 - 4.7\%$ and $3.3 - 4.5\%$, respectively) and data loss ($\epsilon^I = 33\%$ and 37%, respectively).

The 3D fields of mean velocity and turbulence intensity calculated over $T = 750$ s through the first optimal configuration, (i.e. $\Delta\theta = 2.5°$, $\sigma = 1/4$, $m = 5$), are rendered in Figs. 13 and 14, respectively. Furthermore, in Fig. 15, azimuthally-averaged profiles at three downstream locations are also provided for a more insightful comparison. The mean velocity field is reconstructed fairly well regardless of the type of LiDAR scan, due to the careful choice of the fundamental half-wavelength, $\mathbf{\Delta n_0}$,
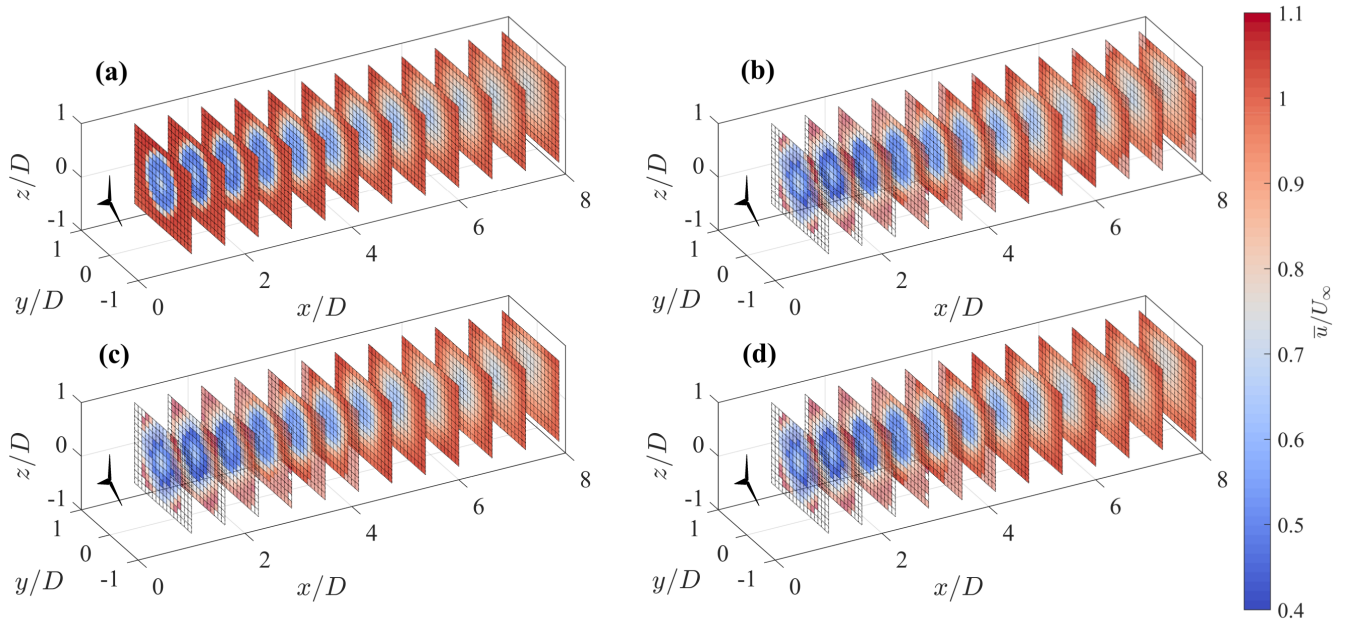
Figure 13. Mean streamwise velocity for $\Delta\theta = 2.5°$, $\sigma = 1/4$, $m = 5$: **(a)** LES; **(b)** ideal; **(c)** step-stare; **(d)** continuous mode LiDAR. The shaded area corresponds to the points rejected after the application of the Petersen-Middleton constraint.
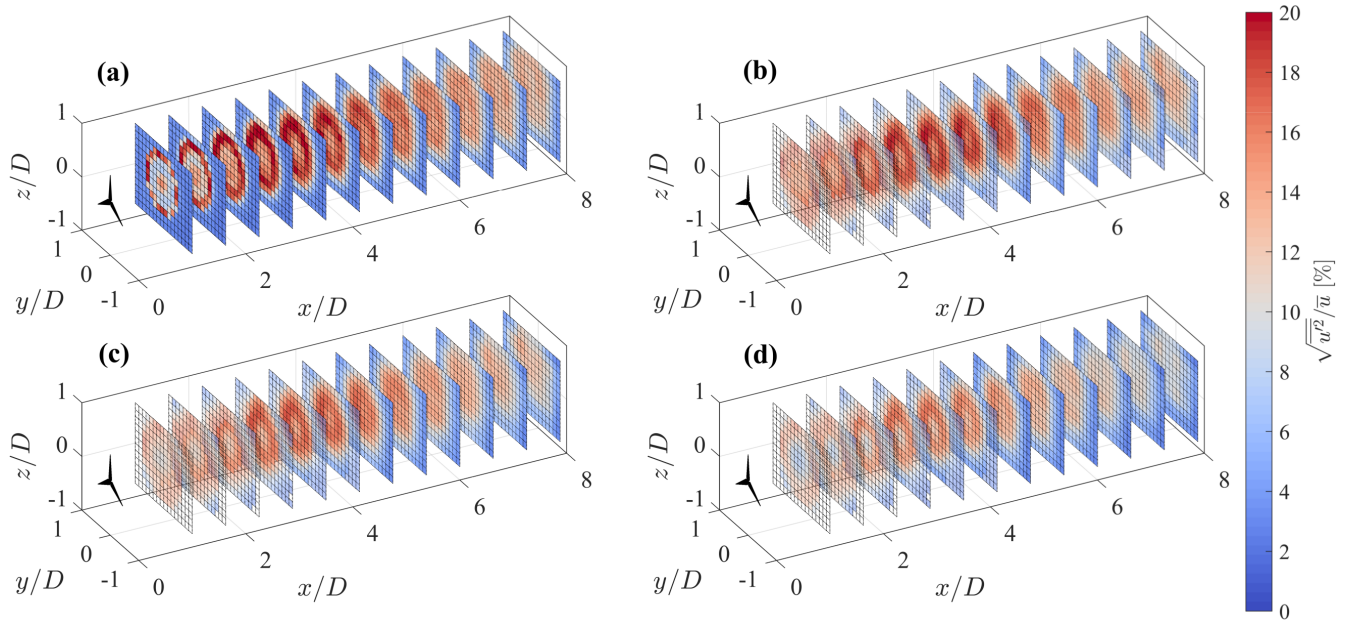
Figure 14. As in Fig. 13 but for streamwise turbulence intensity.
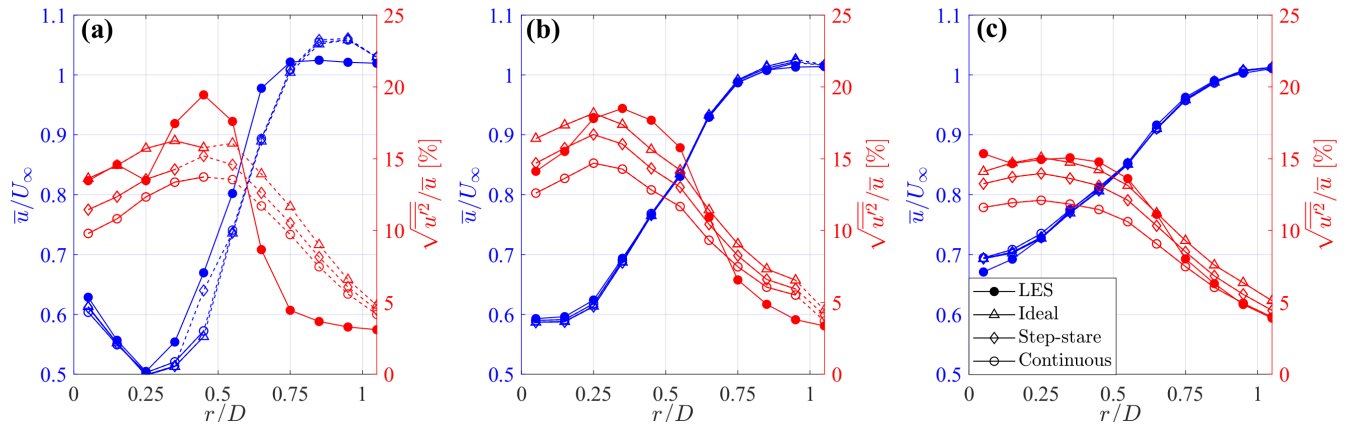
Figure 15. Azimuthally averaged profiles of mean streamwise velocity and turbulence intensity for three downstream locations: **(a)** $x/D = 2.25$; **(b)** $x/D = 4.125$; **(c)** $x/D = 6$. The dashed lines correspond to regions rejected after the application of the Petersen-Middleton constraint.

for this specific flow. On the other hand, the reconstructed turbulence intensity is highly affected by the LiDAR processing, which leads to visible damping of the velocity variance for the step-stare and even more for the continuous mode. The ideal Li-DAR scan, whose acquisition is inherently devoid of any space-time averaging, allows retrieving the correct level of turbulence intensity for locations for $x \geq 4D$, while in the near wake it struggles to recover the thin turbulent ring observed in the wake

575

shear layer. Indeed, such short-wavelength feature has a small response for the chosen settings of the LiSBOA, in particular Δn_0 and σ (see Fig. 3b). On the other hand, any attempt to increase the response of the higher-order moments, for instance by reducing the fundamental half-wavelengths or decreasing the smoothing and the number of iterations, would result in higher data loss and fewer experimental points per grid node.

580   Finally, Figs. 16 and 17 show $\overline{u}/U_\infty$ and $\sqrt{\overline{u'^2}}/\overline{u}$ over several cross-flow planes and for all the combinations of $\sigma - m$ tested for the ideal LiDAR and the optimal angular resolution. For the mean velocity, the most noticeable effect is the increasingly severe data loss consequent to the reduction of σ, which indicates $\sigma = 1/4$ - $m = 5$ as the most effective setting. The turbulence intensity exhibits, in addition to the data loss, a moderate increase in the maximum value for smaller σ, which is due to the higher response of the higher-order statistics (see Table 2). However, this effect is negligible in the far wake were the radial

585 diffusion of the initially sharp turbulent shear layer results in a shift of the energy content towards scales with larger Δn, which are fairly recovered even for $\sigma = 1/4$.

6   Notes on LiSBOA applications

The LiSBOA can be applied to LiDAR datasets that are statistically homogeneous as a function time, t. This statistical property can be ensured with two approaches. The first approach consists in considering LiDAR data collected continuously in time with

590 a given sampling frequency for a period where environmental parameters, such as wind speed and direction, Obukhov length
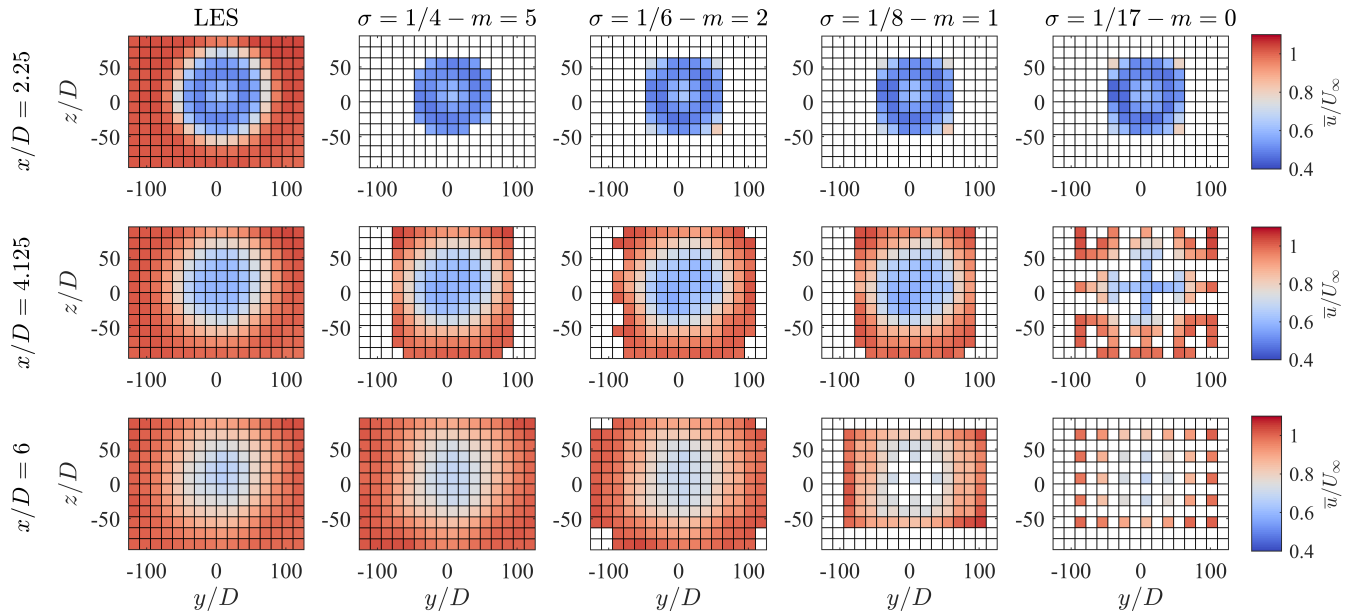
Figure 16. Mean streamwise velocity fields obtained through the ideal LiDAR simulator with $\Delta\theta = 2.5°$ over cross-flow planes at three downstream locations and four combinations of $\sigma - m$, compared with the corresponding LES data.
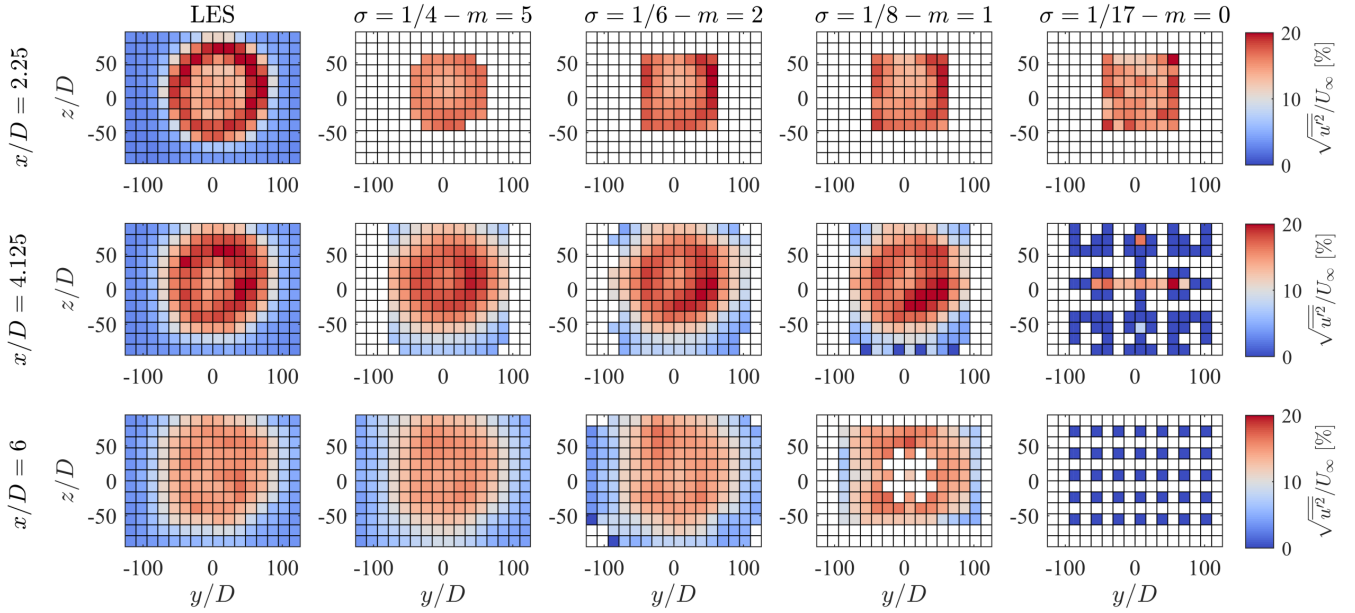
Figure 17. Same as Fig. 16 but for streamwise turbulence intensity.

and Bulk Richardson number for the atmospheric stability regime, are constrained within prefixed intervals (e.g. Banta et al. (2006); Iungo et al. (2013b); Kumer et al. (2015); Puccioni and Iungo (2020)). For instance, the statistical stationarity of a generic flow signal, α, can be verified through the non-stationary index (IST) (Liu et al., 2017):

$$IST = \frac{|\overline{\alpha'\alpha'}_m - \overline{\alpha'\alpha'}|}{\overline{\alpha'\alpha'}}, \tag{19}$$

595 where $\bar{\ }$ represents time-averaging and $\overline{\alpha'\alpha'}_m$ is the mean value of the variance calculated over consecutive non-overlapping sub-periods. The IST values should be lower than a selected threshold depending on the specific flow parameter considered (Foken et al., 2004). A second approach to ensure statistical homogeneity of the LiDAR dataset consists in performing a cluster analysis based on environmental parameters, such as those mentioned above. This can be a fruitful alternative when the application of the first approach leads to too short periods with statistical stationarity and, thus, with low accuracy in the

600 calculation of the turbulent statistics. With the clustering approach, larger datasets can be achieved for each cluster enabling an enhanced statistical accuracy (see e.g. applications of clustering analysis to LiDAR measurements of wind turbine wakes Machefaux et al. (2016a); Bromm et al. (2018); Iungo et al. (2018); Zhan et al. (2019, 2020)).

The results of the LiSBOA for optimal design a wind LiDAR scans are affected by the selection of the input parameters, such as the total sampling time, T, the integral time-scale, τ, the velocity variance, $\overline{u'^2}$, the fundamental half-wavelength, $\boldsymbol{\Delta n_0}$.

605 In this section, we will discuss the sensitivity of the LiSBOA to these input parameters by considering as reference case the volumetric scan performed with the virtual LiDAR technique on the LES dataset analyzed in Sect. 5. The respective results are summarized in Fig. 18.
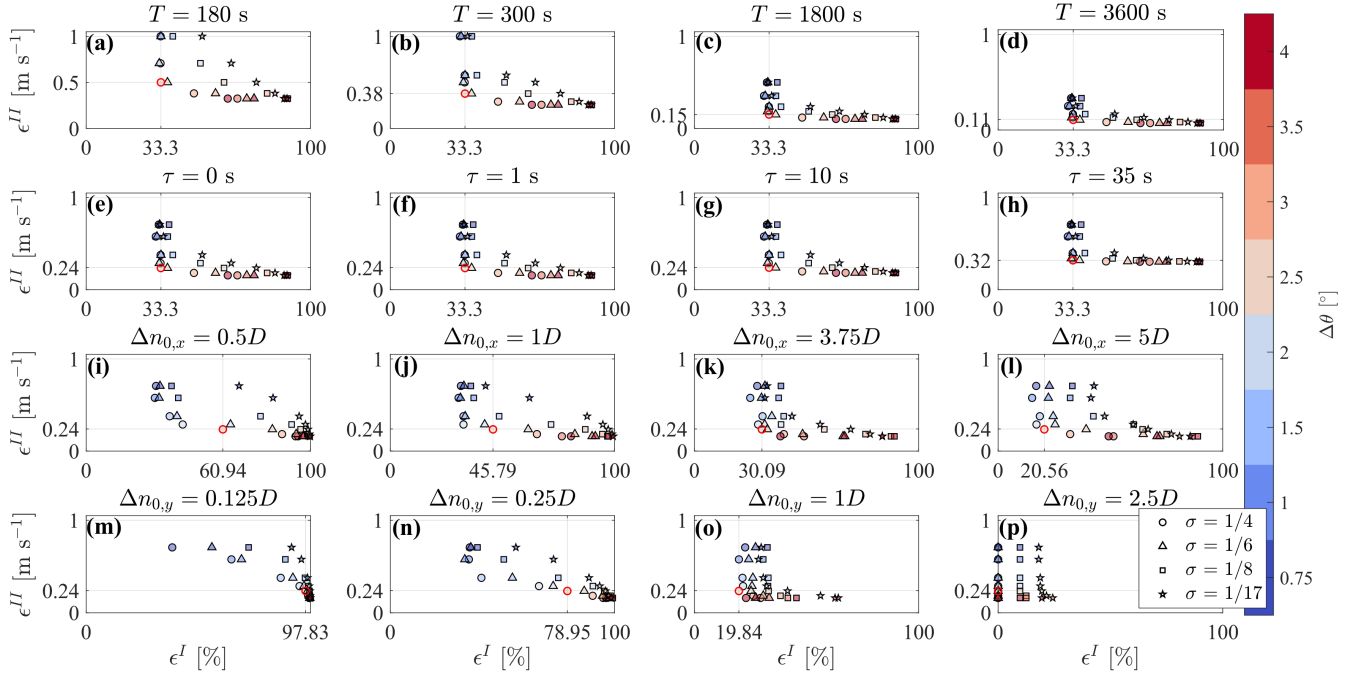
Figure 18. Pareto fronts for the design of the volumetric scan for different inputs: **(a-d)**, sensitivity to total sampling time T; **(e-h)**, sensitivity to integral time-scale, τ; **(i-l)**: sensitivity to streawmise fundamental half-wavelength, $\Delta n_{0,x}$; **(m-p)**: sensitivity to spanwise fundamental half-wavelength, $\Delta n_{0,y}$. The parameters not indicated on the top of the figures are kept equal to the optimal design case identified in Sect. 5.

The total sampling time, T, directly affects the objective function ϵ^{II} through the parameter L, which represents the number of realizations. In Fig. 18a-d, different Pareto fronts are generated for the case under investigation by varying T from 3 minutes to 1 hour. The various Pareto fronts exhibit similar trends for the various values of T, and generally higher values of ϵ^{II}, so lower statistical accuracy, for smaller T. For all the cases, the optimal configuration is still that selected in Sect. 5, namely $\Delta\theta = 2.5° - \sigma = 1/4$.

For this sensitivity study, the characteristic integral time-scale, τ, has been varied between 0 s (completely random uncorrelated data) up to 35 s, the upper value being based on the largest integral length scale in the ABL according to (ESDU, 1975) and considering an advection velocity of 8 m s^{-1}. The respective Pareto fronts reported in Fig. 18e-h show that the optimal LiDAR scan is weakly affected by variations of τ, which is an advantageous feature of the LiSBOA for applications where this parameter cannot be estimated from previous investigations or literature.

Regarding the characteristic velocity variance, $\overline{u'^2}$, it is a multiplicative parameter for the objective function ϵ^{II} (Eq. (14)). Therefore, even though it affects the accuracy in the statistics retrieved, it does not alter the selection of the optimal scanning parameters.

The choice of the fundamental half-wavelength, $\boldsymbol{\Delta n_0}$, deserves special attention since it affects both the optimal scan design and retrieval of data statistics. The fundamental half-wavelength can be considered as the cut-off wavelength of the spatial low-

pass filtering operated by the LiSBOA (Sect. 4). The selection of $\mathbf{\Delta n_0}$ depends mostly on the length of the smallest spatial feature of interest in the flow under investigation, so the Pareto front is likely to be rather sensitive to changes in $\mathbf{\Delta n_0}$. As mentioned in Sect. 4, if the fundamental half-wavelengths are too large compared to the predominant spatial modes, the turbulence statistics may be contaminated by over-smoothing and dispersive stresses, whereas overly small $\mathbf{\Delta n_0}$ may require too small angular and radial resolutions, a longer sampling period, and, thus, a smaller number of repetitions for a given T. Fig. 18i-l shows the Pareto fronts calculated at different $\Delta n_{0,x}$. The previously selected optimal setup ($\Delta\theta = 2.5° - \sigma = 1/4$) still belongs to the optimality frontier. Finally, Fig. 18p-m displays the effect of varying $\Delta n_{0,y}$ on the Pareto front. Unlike the other cases, the shape of the front is very sensitive to $\Delta n_{0,y}$, with a significant increase of data loss, ϵ^I, consequent to refinements of the angular resolution, $\Delta\theta$. The Pareto front correctly indicates that finer angular resolutions are needed to adequately sample a velocity field characterized by smaller wavelengths.

For the sake of completeness, the influence of the different fundamental half-wavelengths on the statistics is assessed by calculating the AE_{95} between LiSBOA and LES for the statistics reconstructed using several values combinations of $\Delta n_{0,x}$ and $\Delta n_{0,y}$ (Fig. 19). Larger values of $\Delta n_{0,x}$ and $\Delta n_{0,y}$ produce detrimental effects on the accuracy for both mean velocity and turbulence intensity due to the over-smoothing of the mean velocity and turbulence intensity field, and dispersive stresses, for the turbulence intensity only. On the other hand, excessively small values of $\Delta n_{0,y}$ exhibit a slight increase in error as a consequence of the smaller number of samples per grid node. Nonetheless, the most relevant effect, in this case, is represented by the high data loss, as already identified in the Pareto front (e.g. Fig. 18m and n). It is noteworthy how the choice of $\mathbf{\Delta n_0} = [2.5, 0.5, 0.5]D$ done in Sect. 5, which was purely based on physical considerations about the expected relevant modes in the near wake, turned out to be the optimal configuration for both $\overline{u}/U_\infty$ and $\sqrt{\overline{u'^2}}/\overline{u}$.

To conclude, we acknowledge that the technical specifications required by LiSBOA (namely τ_a, N_r, Δr) are dependent on the specific LiDAR system used, the contingent atmospheric conditions and the best practices followed by the user. Since these parameters are greatly case-dependent, they will not be discussed further in this context. In general, the selection of the accumulation time and gate length is a trade-off between the need to achieve a target maximum range, while keeping a sufficiently fine radial resolution and sufficient intensity of the back-scattered LiDAR signal. In case of uncertain environmental conditions, it is recommended to check before the deployment the influence of selected combinations of τ_a, N_r, Δr on the Pareto front.

It is noteworthy that the LiSBOA is currently formulated for a single scalar field, namely a velocity component (radial or equivalent horizontal). However, in principle, this procedure can be extended to vector fields, such as fully 3D velocity fields. Furthermore, other constraints can be added for the optimal scanning design, such as imposing a divergence-free velocity field for incompressible flows. Also, some modifications could extend the application of the LiSBOA to other remote sensing instruments, such as sodars and scanning radars.
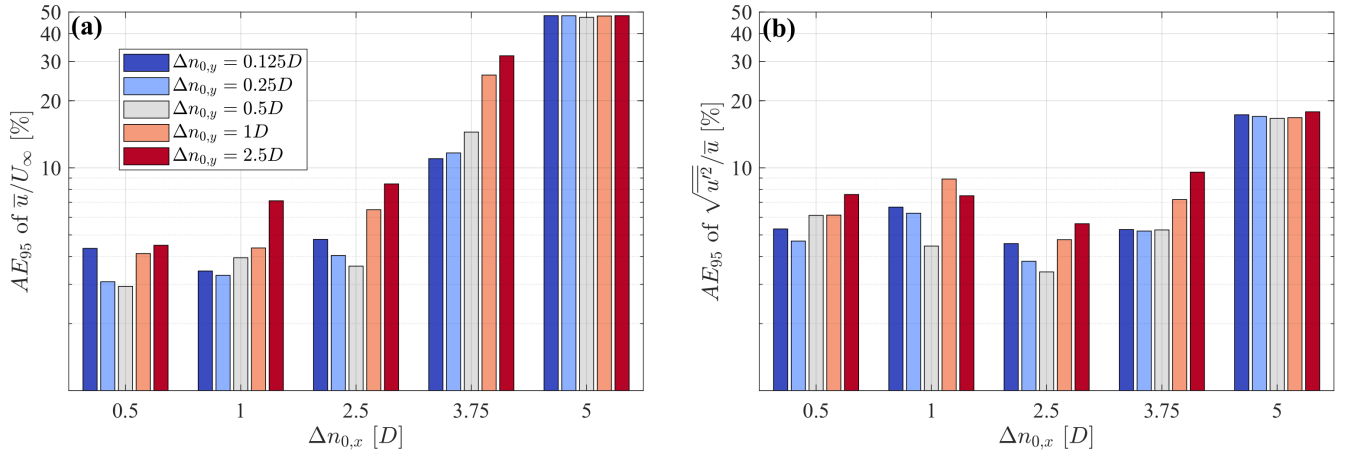
Figure 19. AE_{95} of the statistics reconstructed from virtual LiDAR data for different streamwise and spanwise fundamental half-wavelengths, for the setup $\Delta\theta = 2.5° - \sigma = 1/4$: **(a)** mean streamwise velocity; **b** streamwise turbulence intensity.

7 Conclusions

655 A revisited Barnes objective analysis for sparse, non-uniformly distributed and stationary LiDAR data has been formulated to calculate mean, variance, and higher-order statistics of the wind velocity field over a structured N-dimensional Cartesian grid. This LiDAR Statistical Barnes Objective Analysis (LiSBOA) provides a theoretical framework to quantify the response in the reconstruction of the velocity statistics as a function of the spatial wavelengths of the velocity field under investigation and quantification of the sampling error.

660 The LiSBOA has been validated against volumetric synthetic 3D data generated through Monte-Carlo simulations. The results of this test have shown that the sampling error for a monochromatic scalar field is mainly driven by the data spacing normalized by the half-wavelength.

The LiSBOA framework provides guidelines for the optimal design of scans performed with a scanning Doppler pulsed wind LiDAR and calculation of wind velocity statistics. The optimization problem consists in providing background information
665 about the turbulent flow under investigation, such as total sampling time, expected velocity variance and integral length scales, technical specifications of the LiDAR, such as range gate and accumulation time, and spatial wavelengths of interest for the velocity field. The formulated optimization problem has two cost functions, namely the percentage of grid nodes not satisfying the Petersen-Middleton constraint for the smallest half-wavelength of interest (i.e. lacking adequate spatial resolution to avoid aliasing in the statistics), and the standard deviation of the sample mean. The output of the optimization problem are the LiDAR
670 angular resolution and, for a given response of the mean field, the allowable smoothing parameters and number of iterations to use for the LiSBOA.

The LiSBOA has been validated using a dataset obtained through the virtual LiDAR technique, namely by numerically sampling the turbulent velocity field behind the rotor of a 5 MW turbine obtained from a large eddy simulation (LES). The 3D

mean streamwise velocity and turbulence intensity fields have shown a maximum error with respect to the LES dataset of about
675   4% of the undisturbed wind speed for the mean streawmise velocity and 4% of the streamwise turbulence intensity, in absolute terms. Wake features, such as the high-velocity stream in the hub region and the turbulent shear layer at the wake boundary, have been accurately reconstructed.

In the companion paper (Letizia et al., 2020), the LiSBOA is used to reconstruct the turbulence statistics of utility-scale turbine wakes probed with scanning pulsed Doppler LiDARs. That study also illustrates the detailed preconditioning applied
680   to the raw LiDAR data to extract statistically stationary and normalized radial velocity data and showcases the potential of LiSBOA for wind energy research.

Code availability. The LiSBOA algorithm is implemented in a publicly available code which can be downloaded at the following URL: https://www.utdallas.edu/windflux/.

Appendix A: Derivation of the analytical response function of the LiSBOA

685   The first iteration of the LiSBOA produces a weighted average in space of the scalar field, f, with the weights being Gaussian functions centered at the specific grid nodes, $\boldsymbol{x}$. In the limit of a continuous function defined over an infinite domain, Eq. (3) represents the convolution between the scalar field, f, and the Gaussian weights, w. Therefore, the response function of the LiSBOA, can be expressed in the spectral domain as (Pauley and Wu, 1990):

$$D^0 = \frac{\mathfrak{F}[g^0]}{\mathfrak{F}[f]} = \mathfrak{F}[w], \tag{A1}$$

690   where the operator $\mathfrak{F}$ indicates the Fourier transform (FT). The FT of the weighting function in Eq. (A1) can be conveniently recast as the product of N one-dimensional FT:

$$\mathfrak{F}[w] = \prod_{p=1}^{N} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-x_p^2}{2\sigma^2}} \cdot e^{-\mathrm{i}k_p x_p} dx_p, \tag{A2}$$

where k_p is the directional wavenumber and $\mathrm{i} = \sqrt{-1}$. Hence, by leveraging the closed-form FT of the Gaussian function (Greenberg, 1998):

695    $$\mathfrak{F}\left[\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-x^2}{2\sigma^2}}\right] = e^{\frac{-k^2\sigma^2}{2}} \tag{A3}$$

we get the desired results, i.e.:

$$D^0(\boldsymbol{k}) = e^{-\frac{\sigma^2}{2}|\mathbf{k}|^2}. \tag{A4}$$

Author contributions. SL and GVI developed the LiSBOA and prepared the manuscript. The LiDAR data were generated as a team effort including contributions from all three authors. SL implemented the LiSBOA in a Matlab code under the supervision of GVI.

700 *Competing interests.* No competing interests are present.

Acknowledgements. This research has been funded by a grant from the National Science Foundation CBET Fluid Dynamics, award number 1705837. This material is based upon work supported by the National Science Foundation under grant IIP-1362022 (Collaborative Research: I/UCRC for Wind Energy, Science, Technology, and Research) and from the WindSTAR I/UCRC Members: Aquanis, Inc., EDP Renewables, Bachmann Electronic Corp., GE Energy, Huntsman, Hexion, Leeward Asset Management, LLC, Pattern Energy, EPRI, LM Wind, Texas 705 Wind Tower and TPI Composites. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors. Texas Advanced Computing Center is acknowledged for providing computational resources. The authors thank Stefano Leonardi and Umberto Ciri for sharing the LES dataset.

References

Abkar, M. and Porté-Agel, F.: The effect of free-atmosphere stratification on boundary-layer flow and power output from very large wind farms, Energies, 6, 2338–2361, https://doi.org/10.3390/en6052338, 2013.

Achtemeier, G. L.: The Impact of Data Boundaries upon a Successive Corrections Objective Analysis of Limited-Area Datasets, Mon. Weather Rev., 114, 40–49, https://doi.org/10.1175/1520-0493(1986)114<0040:TIODBU>2.0.CO;2, 1986.

Achtemeier, G. L.: Modification of a Successive Corrections Objective Analysis for Improved Derivative Calculations, Mon. Weather Rev., 117, 78–86, https://doi.org/10.1175/1520-0493(1989)117<0078:MOASCO>2.0.CO;2, 1989.

Aitken, M. L. and Lundquist, J. K.: Utility-Scale Wind Turbine Wake Characterization Using Nacelle-Based Long-Range Scanning Lidar, J. Atmos. Ocean. Tech., 31, 1529–1539, https://doi.org/10.1175/JTECH-D-13-00218.1, 2014.

Arenas, I., García, E., Fu, M. K., Orlandi, P., Hultmark, M., and Leonardi, S.: Comparison between super-hydrophobic, liquid infused and rough surfaces: a direct numerical simulation study, J. Fluid Mech., 869, 500–525, https://doi.org/10.1017/jfm.2019.222, 2019.

Ashton, R., Iungo, G. V., Viola, F., Gallaire, F., and Camarri, S.: Hub vortex instability within wind turbine wakes : Effects of wind turbulence, loading conditions and blade aerodynamics, Physical Review Fluids, 1, 073 603, https://doi.org/10.1103/PhysRevFluids.1.073603, 2016.

Askelson, M. A., Aubagnac, J. P., and Straka, J. M.: An Adaptation of the Barnes Filter Applied to the Objective Analysis of Radar Data, Mon. Weather Rev., 128, 3050–3082, https://doi.org/10.1175/1520-0493(2000)128<3050:AAOTBF>2.0.CO;2, 2000.

Astrup, P., Mikkelsen, T. K., and van Dooren, M. F.: Wind field determination from multiple Spinner-Lidar line-of-sight measurements using linearized CFD, DTU Wind Energy E. 0102, DTU, Department of Wind Energy Frederiksborgvej 399 DK-4000 Roskilde Denmark, 2017.

Banakh, V. A., Smalikho, I. N., Köpp, F., and Werner, C.: Measurements of Turbulent Energy Dissipation Rate with a CW Doppler Lidar in the Atmospheric Boundary Layer, J. Atmos. Ocean. Tech., 16, 1044–1061, https://doi.org/10.1175/1520-0426(1999)016<1044:MOTEDR>2.0.CO;2, 1999.

Banta, R. M., Pichugina, Y. L., and Brewer, W. A.: Turbulent Velocity-Variance Profiles in the Stable Boundary Layer Generated by a Nocturnal Low-Level Jet, J. Atmos. Sci., 63, 2700–2719, https://doi.org/10.1175/JAS3776.1, 2006.

Barnes, S.: Application of Barnes Objective Analysis scheme. Part III: tuning for minimum error., J. Atmos. Ocean. Tech., 11, 1449–1458, https://doi.org/10.1175/1520-0426(1994)011<1459:AOTBOA>2.0.CO;2, 1994c.

Barnes, S. L.: A Technique for Maximizing Details in Numerical Weather Map Analysis, J. Appl. Meteorol., 3, 396–409, https://doi.org/10.1175/1520-0450(1964)003<0396:ATFMDI>2.0.CO;2, 1964.

Barnes, S. L.: Mesoscale Objective Map Analysis Using Weighted Time-Series Observations, NOAA Technical Memorandum ERL NSSL-62, National Severe Storms Laboratory Norman, National Severe Storms Laboratory Norman, OK, 1973.

Barnes, S. L.: Applications of the Barnes Objective Analysis Scheme. Part I: Effects of Undersampling, Wave Position and Station Randomness, J. Atmos. Ocean. Tech., 11, 1433–1448, https://doi.org/10.1175/1520-0426(1994)011<1433:AOTBOA>2.0.CO;2, 1994a.

Bayley, V. G. and Hammersley, J. M.: The "Effective" Number of Independent Observations in an Autocorrelated Time Series, Supplement to the Journal of Royal Statistical Society, 8, 184–197, 1946.

Beck, H. and Kühn, M.: Dynamic Data Filtering of Long-Range Doppler LiDAR Wind Speed Measurements, Remote Sens.-Basel, 9, 561, 2017.

Bell, T. M., Klein, P., Wildmann, N., and Menke, R.: Analysis of flow in complex terrain using multi-Doppler lidar retrievals, Atmos. Meas. Tech., 13, 1357–1371, https://doi.org/10.5194/amt-13-1357-2020, 2020.

Berg, J., Mann, J., Bechmann, A., Courtney, M. S., and Jørgensen, H. E.: The Bolund Experiment. Part I : Flow Over a Steep, Three-dimensional Hill, Bound.-Lay. Meteorol., pp. 219–243, https://doi.org/10.1007/s10546-011-9636-y, 2011.

Bingöl, F.: Complex terrain and wind lidars, DTU PhD Thesis Risø-PhD-52, Technical Univeristy of Denmark, DTU Vindenergi Nils Koppels Allé Bygning 403 2800 Kgs. Lyngby, 2009.

Bingöl, F., Mann, J., and Larsen, G. C.: Light detection and ranging measurements of wake dynamics. Part I: one-dimensional scanning, Wind Energy, 13, 51–61, https://doi.org/10.1002/we.352, 2010.

Bodini, N., Zardi, D., and Lundquist, J. K.: Three-dimensional structure of wind turbine wakes as measured by scanning lidar, Atmos. Meas. Tech., 10, 2881–2896, https://doi.org/10.5194/amt-10-2881-2017, 2017.

Bonin, T. A., Choukulkar, A., Brewer, W. A., Sandberg, S. P., Weickmann, A. M., Pichugina, Y. L., Banta, R. A., Oncley, S. P., and Wolfe, D. E.: Evaluation of turbulence measurement techniques from a single Doppler lidar, Atmos. Meas. Tech., 10, 3021–3039, https://doi.org/10.5194/amt-10-3021-2017, https://www.atmos-meas-tech.net/10/3021/2017/, 2017.

Bradley, E. F.: The influence of thermal stability and angle of incidence on the acceleration of wind up a slope, J. Wind Eng. Ind. Aerod., 15, 231–242, 1983.

Braham, R. R.: Field Experimentation in Weather Modification, J. Am. Stat. Assoc., 74, 57–68, http://www.jstor.org/stable/2286722, 1979.

Bromm, M., Rott, A., Beck, H., Vollmer, L., Steinfeld, G., and Kühn, M.: Field investigation on the influence of yaw misalignment on the propagation of wind turbine wakes, Wind Energy, 21, 1011–1028, https://doi.org/10.1002/we.2210, 2018.

Busch, N. E. and Kristensen, L.: Cup Anemometer Overspeeding, J. of Appl. Meteorol., 15, 1328–1332, https://doi.org/10.1175/1520-0450(1976)015<1328:cao>2.0.co;2, 1976.

Buzzi, A., D Gomis, D., A., P. M., and Alonso, S.: A Method to Reduce the Adverse Impact that Inhomogeneous Station Distributions Have on Spatial Interpolation, Mon. Weather Rev., 119, 2465–2491, https://doi.org/10.1175/1520-0493(1991)119<2465:AMTRTA>2.0.CO;2, 1991.

Caracena, F.: Analytic Approximation of Discrete Field Samples with Weighted Sums and the Gridless Computation of Field Derivatives, J. Atmos. Sci., 44, 3753–3768, https://doi.org/10.1175/1520-0469(1987)044<3753:AAODFS>2.0.CO;2, 1987.

Caracena, F., Barnes, S. L., and Doswell III, C.: Weighting function parameters for objective interpolation of meteorological data, in: Preprints, 10-th Conf. on Weather Forecasting and Analysis, pp. 109–116, 1984.

Carbajo Fuertes, F., Markfort, C. D., and Porté-Agel, F.: Wind Turbine Wake Characterization with Nacelle-Mounted Wind Lidars for Analytical Wake Model Validation, Remote Sens.-Basel, 10, https://doi.org/10.3390/rs10050668, 2018.

Chamorro, L. P. and Porté-Agel, F.: Effects of Thermal Stability and Incoming Boundary-Layer Flow Characteristics on Wind-Turbine Wakes: A Wind-Tunnel Study, Boundary-Layer Meteorology, 136, 515–533, https://doi.org/10.1007/s10546-010-9512-1, 2010.

Choukulkar, A., Brewer, W. A., Sandberg, S. P., Weickmann, A., Bonin, T. A., Hardesty, R. M., Lundquist, J. K., Delgado, R., Iungo, G. V., Ashton, R., Debnath, M., Bianco, L., Wilczak, J. M., Oncley, S., and Wolfe, D.: Evaluation of single and multiple Doppler lidar techniques to measure complex flow during the XPIA field campaign, Atmos. Meas. Tech., 10, 247–264, https://doi.org/10.5194/amt-10-247-2017, 2017.

Ciri, U., Rotea, M., Santoni, C., and Leonardi, S.: Large-eddy simulations with extremum-seeking control for individual wind turbine power optimization, Wind Energy, 20, 1617–1634, https://doi.org/10.1002/we.2112, 2017.

Clive, P. J. M., Dinwoodie, I., and Quail, F.: Direct measurement of wind turbine wakes using Remote Sens.-Basel, Proceedings of the EWEA, 2011.

Cressman, G. P.: An Operational Objective Analysis System, Mon. Weather Rev., 87, 367–374, https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2, 1959.

Cuerva, A. and Sanz-Andrés, A.: On sonic anemometer measurement theory, J. Wind Eng. Ind. Aerod., 88, 25–55, https://doi.org/10.1016/S0167-6105(00)00023-4, 2000.

785 Cushman-Roisin, B. and Beckers, J. M.: Introduction, in: Introduction to Geophysical Fluid Dynamics, pp. 7–11, Academic Press, 1990a.

Cushman-Roisin, B. and Beckers, J. M.: Stratification effects, in: Introduction to Geophysical Fluid Dynamics, p. 319, Academic Press, 1990b.

Davies, F., Middleton, D. R., and Bozier, K. E.: Urban air pollution modelling and measurements of boundary layer height, Atmos. Environ., 41, 4040–4049, https://doi.org/10.1016/j.atmosenv.2007.01.015, 2007.

790 Debnath, M., Iungo, G. V., Ashton, R., Alan Brewer, W., Choukulkar, A., Delgado, R., Lundquist, J. K., Shaw, W. J., Wilczak, J. M., and Wolfe, D.: Vertical profiles of the 3-D wind velocity retrieved from multiple wind lidars performing triple range-height-indicator scans, Atmos. Meas. Tech., 10, 431–444, https://doi.org/10.5194/amt-10-431-2017, 2017a.

Debnath, M., Iungo, G. V., Brewer, W. A., Choukulkar, A., Delgado, R., Gunter, S., Lundquist, J. K., Schroeder, J. L., Wilczak, J. M., and Wolfe, D.: Assessment of virtual towers performed with scanning wind lidars and Ka-band radars during the XPIA experiment, Atmos. 795 Meas. Tech., pp. 1215–1227, https://doi.org/10.5194/amt-10-1215-2017, 2017b.

Debnath, M., Santoni, C., Leonardi, S., and Iungo, G. V.: Towards reduced order modelling for predicting the dynamics of coherent vorticity structures within wind turbine wakes, Philos. T. R. Soc. A., 375, https://doi.org/10.1098/rsta.2016.0108, 2017c.

Doswell, C. A.: Obtaining Meteorologically Significant Surface Divergence Fields Through the Filtering Property of Objective Analysis, Mon. Weather Rev., 105, 885–892, https://doi.org/10.1175/1520-0493(1977)105<0885:OMSSDF>2.0.CO;2, 1977.

800 Duncan, Jr., J. B., Hirth, B. D., and Schroeder, J. L.: Enhanced estimation of boundary layer advective properties to improve space-to-time conversion processes for wind energy applications, Wind Energy, 22, 1203–1218, https://doi.org/10.1002/we.2350, https://onlinelibrary.wiley.com/doi/abs/10.1002/we.2350, 2019.

Eberhard, W. L., Cupp, R. E., and Healy, K. R.: Doppler Lidar Measurement of Profiles of Turbulence and Momentum Flux, J. Atmos. Ocean. Tech., 6, 809–819, https://doi.org/10.1175/1520-0426(1989)006<0809:DLMOPO>2.0.CO;2, 1989.

805 El-Asha, S., Zhan, L., and Iungo, G. V.: Quantification of power losses due to wind turbine wake interactions through SCADA, meteorological and wind LiDAR data, Wind Energy, 20, 1823–1839, https://doi.org/10.1002/we.2123, 2017.

Emeis, S.: Surface-Based Remote Sensing of the Atmospheric Boundary Layer, Springer: Berlin/Heidelberg, 2010.

Emeis, S., Frank, H. R., and Fiedler, F.: Modification of air flow over an escarpment - Results from the Hjardemal experiment, Bound.-Lay. Meteorol., 74, 131–161, 1995.

810 Endlich, R. M. and Mancuso, R. L.: Objective analysis of environmental conditions associated with severe thunderstorms and tornadoes, Mon. Weather Rev., 96, 342–350, https://doi.org/10.1126/science.27.693.594, 1968.

ESDU: Characteristics of atmospheric turbulence near the ground. Part III: Variations in space and time for strong winds (neutral atmosphere), Engineering Sciences Data Unit, 1975.

Fernando, H. J. S., Mann, J., Palma, J. M. L. M., Lundquist, J. K., Barthelmie, R. J., Belo-Pereira, M., Brown, W., Chow, F. K., and Gerz, 815 T., e. a.: Peering into microscale details of mountain winds, B. Am. Meteorol. Soc., 100, 799–820, https://doi.org/10.1175/BAMS-D-17-0227.1, 2019.

Foken, T., Göockede, M., Mauder, M., Mahrt, L., B., A., and W., M.: Post-Field Data Quality Control, in: Handbook of Micrometeorology, Atmospheric and Oceanographic Sciences Library, vol 29. Springer, Dordrecht, 2004.

Frehlich, R.: Effects of Wind Turbulence on Coherent Doppler Lidar Performance, J. Atmos. Ocean. Tech., 14, 54–75, https://doi.org/10.1175/1520-0426(1997)014<0054:EOWTOC>2.0.CO;2, 1997.

Frehlich, R. and Cornman, L.: Estimating Spatial Velocity Statistics with Coherent Doppler Lidar, J. Atmos. Ocean. Tech., 19, 355–366, https://doi.org/10.1175/1520-0426-19.3.355, 2002.

Frisch, A. S.: On the measurement of second moments of turbulent wind velocity with a single Doppler radar over non-homogeneous terrain, Bound.-Lay. Meteorol., 54, 29–39, 1991.

Fuertes Carbajo, F. and Porté-Agel, F.: Using a Virtual Lidar Approach to Assess the Accuracy of the Volumetric Reconstruction of a Wind Turbine Wake, Remote Sens.-Basel, 10, https://doi.org/10.3390/rs10050721, 2018.

Gal-Chen, T., Xu, M., and Eberhard, W. L.: Estimations of atmospheric boundary layer fluxes and other turbulence parameters from Doppler lidar data, J. Geophys. Res.-Atmos., 97, 18 409–18 423, https://doi.org/10.1029/91JD03174, 1992.

Garcia, E. T., Aubrun, S., Boquet, M., Royer, P., Coupiac, O., and Girard, N.: Wake meandering and its relationship with the incoming wind characteristics: a statistical approach applied to long-term on-field observations, Journal of Physics: Conference Series, 854, 012 045, https://doi.org/10.1088/1742-6596/854/1/012045, 2017.

George, W., Beuther, P. D., and Lumley, J. L.: Processing of Random Signals, in: Proceedings of the Dynamic Flow Conference, pp. 757–800, 1978.

Gomis, D. and Alonso, S.: Diagnosis of a Cyclogenetic Event in the Western Mediterranean Using an Objective Technique for Scale Separation, Mon. Weather Rev., 118, 723–736, https://doi.org/10.1175/1520-0493(1990)118<0723:DOACEI>2.0.CO;2, 1990.

Greenberg, M.: Advanced Engineering Mathematics, 2nd edition, Prentice Hall, 1998.

Halios, C. H. and Barlow, J. F.: Observations of the Morning Development of the Urban Boundary Layer Over London, UK, Taken During the ACTUAL Project, Bound.-Lay. Meteorol., 166, 395–422, https://doi.org/10.1007/s10546-017-0300-z, 2018.

Haugen, D. A., Kaimal, J. C., and Bradley, E. F.: An experimental study of Reynolds stress and heat flux in the atmospheric surface layer, Q. J. Roy. Meteorol. Soc., 97, 168–180, https://doi.org/10.1002/qj.49709741204, 1971.

Holzäpfel, F., Stephan, A., Heel, T., and Körner, S.: Enhanced wake vortex decay in ground proximity triggered by plate lines, Aircr. Eng.g Aerosp.Tech., 88, 206–214, https://doi.org/10.1108/AEAT-02-2015-0045, 2016.

Huang, M., Gao, Z., Miao, S., Chen, F., LeMone, M. A., Li, J., Hu, F., and Wang, L.: Estimate of Boundary-Layer Depth Over Beijing, China, Using Doppler Lidar Data During SURF-2015, Bound.-Lay. Meteorol., 162, 503–522, https://doi.org/10.1007/s10546-016-0205-2, 2017.

Iungo, G. V. and Porté-Agel, F.: Volumetric Lidar Scanning of Wind Turbine Wakes under Convective and Neutral Atmospheric Stability Regimes, J. Atmos. Ocean. Tech., 31, 2035–2048, https://doi.org/10.1175/JTECH-D-13-00252.1, 2014.

Iungo, G. V., Viola, F., Camarri, S., Porté-Agel, F., and Gallaire, F.: Linear stability analysis of wind turbine wakes performed on wind tunnel measurements, J. Fluid Mech., 737, 499–526, https://doi.org/10.1017/jfm.2013.569, http://www.journals.cambridge.org/abstract{_}S0022112013005697, 2013a.

Iungo, G. V., Wu, Y., and Porté-Agel, F.: Field Measurements of Wind Turbine Wakes with Lidars, J. Atmos. Ocean. Tech., 30, 274–287, https://doi.org/10.1175/JTECH-D-12-00051.1, 2013b.

Iungo, G. V., Letizia, S., and Zhan, L.: Quantification of the axial induction exerted by utility-scale wind turbines by coupling LiDAR measurements and RANS simulations, Journal of Physics: Conference Series, 1037, https://doi.org/10.1088/1742-6596/1037/7/072023, 2018.

Ivanell, S., Mikkelsen, R., Sørensen, J. N., and Henningson, D.: Stability analysis of the tip vortices of a wind turbine, Wind Energy, 13, 705–715, https://doi.org/10.1002/we, http://onlinelibrary.wiley.com/doi/10.1002/we.1608/full, 2010.

Jonkman, J., Butterfield, S., Musial, W., and Scott, G.: Definition of a 5-MW Reference Wind Turbine for Offshore System Development, Technical report NREL/TP-500-38060, NREL, 1617 Cole Boulevard, Golden, CO, 2009.

Käsler, Y., S., R., Simmet, R., and Kühn, M.: Wake Measurements of a Multi-MW Wind Turbine with Coherent Long-Range Pulsed Doppler Wind Lidar, J. Atmos. Ocean. Tech., 27, 1529–1532, https://doi.org/10.1175/2010JTECHA1483.1, 2010.

Kim, D., Kim, T., Oh, G., Huh, J., and Ko, K.: A comparison of ground-based LiDAR and met mast wind measurements for wind resource assessment over various terrain conditions, J. Wind Eng. Ind. Aerod., 158, 109–121, https://doi.org/10.1016/j.jweia.2016.09.011, http://dx.doi.org/10.1016/j.jweia.2016.09.011, 2016.

Koch, S. E., Des Jardins, M., and Kocin, P. J.: An Interactive Barnes objective Map Analysis Scheme for Use with Satellite and Conventional Data, J. Clim. Appl. Meteorol., 22, 1487–1503, https://doi.org/10.1175/1520-0450(1983)022<1487:AIBOMA>2.0.CO;2, 1983.

Kongara, S., Calhoun, R., Choukulkar, A., and Boldi, M. O.: Velocity retrieval for coherent Doppler lidar, Int. J. Remote Sens., 33, 3596–3613, https://doi.org/10.1080/01431161.2011.631948, 2012.

Köpp, F., Rahm, S., Smalikho, I., Dolfi, A., Cariou, J. P., Harris, M., and Young, R. I.: Comparison of wake-vortex parameters measured by pulsed and continuous-wave lidars, J. Aircraft, 42, 916–923, https://doi.org/10.2514/1.8177, 2005.

Krishnamurthy, R., Choukulkar, A., Calhoun, R., Fine, J., Oliver, A., and Barr, K. S.: Coherent Doppler lidar for wind farm characterization, Wind Energy, 16, 189–206, https://doi.org/10.1002/we, 2013.

Kropfli, R. A.: Single Doppler Radar Measurements of Turbulence Profiles in the Convective Boundary Layer, J. Atmos. Ocean. Tech., 3, 305–314, https://doi.org/10.1175/1520-0426(1986)003<0305:SDRMOT>2.0.CO;2, 1986.

Kumer, V. M., Reudera, J., Svardalc, B., Sætrec, C., and Eecen, P.: Characterisation of Single Wind Turbine Wakes with Static and Scanning WINTWEX-W LiDAR Data, Energy Proced., 80, 245 – 254, https://doi.org/https://doi.org/10.1016/j.egypro.2015.11.428, 12th Deep Sea Offshore Wind R and D Conference, EERA DeepWind 2015, 2015.

Kunkel, G. J. and Marusic, I.: Study of the near-wall-turbulent region of the high-Reynolds-number boundary layer using an atmospheric flow, J. Fluid Mech., 548, 375–402, https://doi.org/10.1017/S0022112005007780, 2006.

Letizia, S., Zhan, L., and Iungo, G. V.: LiSBOA: LiDAR Statistical Barnes Objective Analysis for optimal design of LiDAR scans and retrieval of wind statistics. Part II: Applications to synthetic and real LiDAR data of wind turbine wakes, Atmos. Meas. Tech. Discussions, 2020, 1–40, https://doi.org/10.5194/amt-2020-228, https://amt.copernicus.org/preprints/amt-2020-228/, 2020.

Lhermitte, R. M.: Note on the Observation of Small-Scale Atmospheric Turbulence by Doppler Radar Techniques, Radio Sci., 4, 1241–1246, https://doi.org/10.1029/RS004i012p01241, 1969.

Liu, H. Y., , Bo, T. L., and Liang, Y. R.: The variation of large-scale structure inclination angles in high Reynolds number atmospheric surface layers, Physics of Fluids, 29, https://doi.org/10.1063/1.4978803, http://dx.doi.org/10.1063/1.4978803, 2017.

Liu, Z., Barlow, J. F., Chan, P. W., Fung, J. C. H., Li, Y., Ren, C., Mak, H. W. L., and Ng, E.: A review of progress and applications of pulsed Doppler Wind LiDARs, Remote Sens.-Basel, 11, 1–47, https://doi.org/10.3390/rs11212522, 2019.

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

Lundquist, J. K., Wilczak, J. M., Ashton, R., Bianco, L., Brewer, W. A., Choukulkar, A., Clifton, A., Debnath, M., Delgado, R., Friedrich, K., Gunter, S., Hamidi, A., Iungo, G. V., Kaushik, A., Kosovic, B., Langan, P., Lass, A., Lavin, E., Lee, J. C. Y., McCaffrey, K. L., Newsom, R. K., Noone, D. C., Oncley, S. P, ., Quelet, P. T., Sandberg, S. P., Schroeder, J. L., Shaw, W. J., Sparling, L., Martin, C. S., Pe, A. S., Strobach, E., Tay, K., Vanderwende, B. J., Weickmann, A., Wolfe, D., and Worsnop, R.: Assessing state-of-The-Art capabilities for

probing the atmospheric boundary layer the XPIA field campaign, B. Am. Meteorol. Soc., 98, 289–314, https://doi.org/10.1175/BAMS-D-15-00151.1, 2017.

Machefaux, E., Larsen, G. C., Koblitz, T., Troldborg, N., Kelly, M. C., Chougule, A., Hansen, K. S., and Rodrigo, J. S.: An experimental and numerical study of the atmospheric stability impact on wind turbine wakes, Wind Energy, 19, 1785–1805, https://doi.org/10.1002/we.1950, 2016a.

Machefaux, E., Larsen, G. C., Troldborg, N., Hansen, K. S., Angelou, N., Mikkelsen, T., and Mann, J.: Investigation of wake interaction using full-scale lidar measurements and large eddy simulation, Wind Energy, 19, 1535–1551, https://doi.org/10.1002/we.1936, 2016b.

Maddox, R. A.: An Objective Technique for Separating Macroscale and Mesoscale Features in Meteorological Data, Mon. Weather Rev., 108, 1108–1121, https://doi.org/10.1175/1520-0493(1980)108<1108:AOTFSM>2.0.CO;2, 1980.

Mann, J., Cariou, J. P., Courtney, M., Parmentier, R., Mikkelsen, T., Wagner, R., Lindelöw, P., Sjöholm, M., and Enevoldsen, K.: Comparison of 3D turbulence measurements using three staring wind lidars and a sonic anemometer, https://doi.org/10.1127/0941-2948/2009/0370, 2009.

Mann, J., Peña, A., Bingöl, F., Wagner, R., and Courtney, M. S.: Lidar Scanning of Momentum Flux in and above the Atmospheric Surface Layer, J. Atmos. Ocean. Tech., 27, 959–976, https://doi.org/10.1175/2010JTECHA1389.1, 2010.

Mann, J., Menke, R., Vasiljević, N., Berg, J., and Troldborg, N.: Challenges in using scanning lidars to estimate wind resources in complex terrain, Journal of Physics: Conference Series, 1037, 0–8, https://doi.org/10.1088/1742-6596/1037/7/072017, 2018.

Manninen, A. J., O'Connor, E. J., Vakkari, V., and Petäjä, T.: A generalised background correction algorithm for a Halo Doppler lidar and its application to data from Finland, Atmos. Meas. Tech., 9, 817–827, https://doi.org/10.5194/amt-9-817-2016, 2016.

Mayor, S. D., Lenschow, D. H., Schwiesow, R. L., Mann, J., Frush, C. L., and Simon, M. K.: Validation of NCAR 10.6-μm CO_2 Doppler Lidar Radial Velocity Measurements and Comparison with a 915-MHz Profiler, J. Atmos. Ocean. Tech., 14, 1110–1126, https://doi.org/10.1175/1520-0426(1997)014<1110:VONMCD>2.0.CO;2, 1997.

Mirocha, J. D., Rajewski, D. A., Marjanovic, N., Lundquist, J. K., Kosović, B., Draxl, C., and Churchfield, M. J.: Investigating wind turbine impacts on near-wake flow using profiling lidar data and large-eddy simulations with an actuator disk model, J. Renew. Sustain. Ener., 7, 1–21, https://doi.org/10.1063/1.4928873, 2015.

Mortensen, N. G.: Wind measurements for wind energy applications—a review, Proceedings of the 16th British Wind Energy Association Conference', Stirling, Scotland, pp. 353–360, 1994.

Muñoz-Esparza, D., Cañadillas, B., Neumann, T., and Van Beeck, J.: Turbulent fluxes, stability and shear in the offshore environment: Mesoscale modelling and field observations at FINO1, J. Renew. Sustain. Ener., 4, https://doi.org/10.1063/1.4769201, 2012.

Newsom, R., Berg, L., Shaw, W. J., and Fischer, M. L.: Turbine-scale wind field measurements using dual-Doppler lidar, Wind Energy, 18, 219–235, https://doi.org/10.1002/we.1691, 2014.

Newsom, R. K. and Banta, R. M.: Assimilating Coherent Doppler Lidar Measurements into a Model of the Atmospheric Boundary Layer. Part I: Algorithm Development and Sensitivity to Measurement Error, J. Atmos. Ocean. Tech., 21, 1328–1345, https://doi.org/10.1175/1520-0426(2004)021<1328:ACDLMI>2.0.CO;2, 2004.

Newsom, R. K., Calhoun, R., Ligon, D., and Allwine, J.: Linearly Organized Turbulence Structures Observed Over a Suburban Area by Dual-Doppler Lidar, Bound.-Lay. Meteorol., 127, 111–130, https://doi.org/10.1007/s10546-007-9243-0, 2008.

Newsom, R. K., Alan Brewer, W., Wilczak, J. M., Wolfe, D. E., Oncley, S. P., and Lundquist, J. K.: Validating precision estimates in horizontal wind measurements from a Doppler lidar, Atmos. Meas. Tech., 10, 1229–1240, https://doi.org/10.5194/amt-10-1229-2017, 2017.

O'Connor, E. J., Illingworth, A. J., Brooks, I. M., Westbrook, C. D., Hogan, R. J., Davies, F., and Brooks, B. J.: A Method for Estimating the Turbulent Kinetic Energy Dissipation Rate from a Vertically Pointing Doppler Lidar and Independent Evaluation from Balloon-Borne In Situ Measurements, J. Atmos. Ocean. Tech., 27, 1652–1664, https://doi.org/10.1175/2010JTECHA1455.1, 2010.

Orlanski, I.: A simple boundary condition for unbounded hyperbolic flows, J. of Comput. Phys., 21, 251 – 269, https://doi.org/10.1016/0021-9991(76)90023-1, 1976.

Pardyjak, E. R. and Stoll, R.: Improving measurement technology for the design of sustainable cities, Meas. Sci. Technol., 28, https://doi.org/10.1088/1361-6501/aa7c77, 2017.

Pashow, M., Parlange, M. B., and Porté-Agel, F.: On Monin–Obukhov similarity in the stable atmospheric boundary layer, Bound.-Lay. Meteorol., 99, 225–248, https://doi.org/10.1023/A, 2001.

Pauley, P. M. and Wu, X.: The Theoretical, Discrete and Actual Response of the Barnes Objective Analysis Scheme for One- and Two-Dimensional Fields, Mon. Weather Rev., 118, 1145–1164, https://doi.org/10.1175/1520-0493(1990)118<1145:TTDAAR>2.0.CO;2, 1990.

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

Petersen, D. P. and Middleton, D.: Sampling and reconstruction of wave-number-limited functions in N-dimensional Euclidean spaces, Inform. Control., 5, 279–323, https://doi.org/10.1016/S0019-9958(62)90633-2, 1962.

Puccioni, M. and Iungo, G. V.: Spectral correction of turbulent energy damping on wind LiDAR measurements due to range-gate averaging, Atmos. Meas. Tech. Discussions, pp. 1–28, 2020.

Rajewski, D. A., Takle, E. S., Lundquist, J. K., Oncley, S., Prueger, J. H., Horst, T. W., Rhodes, M. E., Pfeiffer, R., Hatfield, J. L., Spoth, K. K., and Doorenbos, R. K.: Crop wind energy experiment (CWEX): Observations of surface-layer, boundary layer and mesoscale interactions with a wind farm, B. Am. Meteorol. Soc., 94, 655–672, https://doi.org/10.1175/BAMS-D-11-00240.1, 2013.

Risan, A., Lund, J. A., Chang, C. Y., and Sætran, L.: Wind in Complex Terrain - Lidar Measurements for Evaluation of CFD Simulations, Remote Sens.-Basel, 10, 59, 2018.

Rye, B. and Hardesty, M.: Discrete Spectral Peak Estimation in Incoherent Backscatter Heterodyne Lidar. I: Spectral Accumulation and the Cramer-Rao Lower Bound, IEEE T. Geosci. Remote, 31, 16–27, https://doi.org/10.1109/36.210440, 1993.

Santoni, C., Ciri, U., Rotea, M., and Leonardi, S.: Development of a high fidelity CFD code for wind farm control, Proceedings of the American Control Conference, 2015-July, 1715–1720, https://doi.org/10.1109/ACC.2015.7170980, 2015.

Santoni, C., Carrasquillo, K., Arenas-Navarro, I., and Leonardi, S.: Effect of tower and nacelle on the flow past a wind turbine, Wind Energy, 20, 1927–1939, https://doi.org/10.1002/we.2130, 2017.

Sasaki, Y.: A theoretical interpretation of anisotropically weighted smoothing on the basis of numerical variational analysis, Mon. Weather Rev., 99, 698–707, 1971.

Sathe, A. and Mann, J.: A review of turbulence measurements using ground-based wind lidars, Atmos. Meas. Tech., 6, 3147–3167, https://doi.org/10.5194/amt-6-3147-2013, https://www.atmos-meas-tech.net/6/3147/2013/, 2013.

Sathe, A., Mann, J., Gottschall, J., and Courtney, M. S.: Can Wind Lidars Measure Turbulence?, J. Atmos. Ocean. Tech., 28, 853–868, https://doi.org/10.1175/JTECH-D-10-05004.1, 2011.

Schween, J. H., Hirsikko, A., Löhnert, U., and Crewell, S.: Mixing-layer height retrieval with ceilometer and Doppler lidar: From case studies to long-term assessment, Atmos. Meas. Tech., 7, 3685–3704, https://doi.org/10.5194/amt-7-3685-2014, 2014.

Seaman, R. S.: Tuning the Barnes Objective Analysis Parameters by Statistical Interpolation Theory, J. Atmos. Ocean. Tech., 6, 993–1000, https://doi.org/10.1175/1520-0426(1989)006<0993:TTBOAP>2.0.CO;2, 1989.

970 Sekar, A. P. K., van Dooren, M. F., Mikkelsen, T., Sjöholm, M., Astrup, P., and Kühn, M.: Evaluation of the LINCOM wind field reconstruction method with simulations and full-scale measurements, Journal of Physics: Conference Series, 1037, 052 008, https://doi.org/10.1088/1742-6596/1037/5/052008, 2018.

Shannon, C. E.: Communication in the Presence of Noise, Proceedings of the IEEE, 72, 1192–1201, https://doi.org/10.1109/PROC.1984.12998, 1984.

975 Smalikho, I. N.: On measurement of dissipation rate of the turbulent energy with a CW lidar, Atmos. Ocean. Optics, 8, 788–793, 1995.

Smith, D. R. and Leslie, F. W.: Error Determination of a Successive Correction Type Objective Analysis Scheme, J. Atmos. Ocean. Tech., 1, 120–130, https://doi.org/10.1175/1520-0426(1984)001<0120:EDOASC>2.0.CO;2, 1984.

Smith, D. R., Pumphry, M. E., and Snow, J. T.: A Comparison of Errors in objectively Analyzed Fields for Uniform and Nonuniform Station Distributions, J. Atmos. Ocean. Tech., 3, 84–97, https://doi.org/10.1175/1520-0426(1986)003<0084:ACOEIO>2.0.CO;2, 1986.

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

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

985 Stull, R. B.: Preface, in: An introduction to Boundary Layer Meteorology, p. xi, Kluwer Academic Publisher, P.O. 80x 17. 3300 AA Dordredt. The NetherlandS, 1988.

Tang, W., Chan, P. W., and Haller, G.: Lagrangian coherent structure analysis of terminal winds detected by lidar. Part I: Turbulence structures, J. Appl. Meteorol. Clim., 50, 325–338, https://doi.org/10.1175/2010JAMC2508.1, 2011.

Taylor, P. A. and Teunissen, H. W.: The Askervin hill project: overview and background data, Bound.-Lay. Meteorol., 39, 15–39, 1987.

990 Thobois, L., Cariou, J. P., and Gultepe, I.: Review of Lidar-Based Applications for Aviation Weather, Pure Appl. Geophys., 176, 1959–1976, https://doi.org/10.1007/s00024-018-2058-8, 2019.

Trapp, R. J. and Doswell, C. A.: Radar data objective analysis, J. Atmos. Ocean. Tech., 17, 105–120, https://doi.org/10.1175/1520-0426(2000)017<0105:RDOA>2.0.CO;2, 2000.

Trujillo, J. J., F., B., Larsen, G. C., Mann, J., and Kühn, M.: Light detection and ranging measurements of wake dynamics. Part II: two-995 dimensional scanning, Wind Energy, 14, 61–75, https://doi.org/10.1002/we.402, 2011.

Trujillo, J. J., Seifert, J. K., Würth, I., Schlipf, D., and Kühn, M.: Full-field assessment of wind turbine near-wake deviation in relation to yaw misalignment, Wind Energy Science, 1, 41–53, https://doi.org/10.5194/wes-1-41-2016, 2016.

Vakkari, V., O'Connor, E. J., Nisantzi, A., Mamouri, R. E., and Hadjimitsis, D. G.: Low-level mixing height detection in coastal locations with a scanning Doppler lidar, Atmos. Meas. Tech., 8, 1875–1885, https://doi.org/10.5194/amt-8-1875-2015, 2015.

1000 Vakkari, V., Manninen, A. J., O'Connor, E. J., Schween, J. H., Van Zyl, P. G., and Marinou, E.: A novel post-processing algorithm for Halo Doppler lidars, Atmos. Mea. Tech., 12, 839–852, https://doi.org/10.5194/amt-12-839-2019, 2019.

Van Dooren, M. F., Trabucchi, D., and Kühn, M.: A Methodology for the Reconstruction of 2D Horizontal Wind Fields of Wind Turbine Wakes Based on Dual-Doppler Lidar Measurements, Remote Sens.-Basel, 8, https://doi.org/10.3390/rs8100809, 2016.

Viola, F., Iungo, G. V., Camarri, S., Porté-Agel, F., and Gallaire, F.: Prediction of the hub vortex instability in a wind turbine wake: Stability 1005 analysis with eddy-viscosity models calibrated on wind tunnel data, J. Fluid Mech., 750, R1, https://doi.org/10.1017/jfm.2014.263, 2014.

Wang, H. and Barthelmie, R. J.: Wind turbine wake detection with a single Doppler wind lidar, Journal of Physics: Conference Series, 625, 012 017, https://doi.org/10.1088/1742-6596/625/1/012017, 2015.

Wheeler, A. J. and Ganji, A. R.: Measuring Fluid Flow Rate, Fluid Velocity, Fluid Level and Combustion Pollutants, in: Introduction to engineering experimentation, Pearson Higher Education, 1 Lake St., Upper Saddle River, New Jersey, 07458., 2010a.

Wheeler, A. J. and Ganji, A. R.: Experimental Uncertainty Analysis, in: Introduction to Engineering Experimentation, Pearson Higher Education, 1 Lake St., Upper Saddle River, New Jersey, 07458., 2010b.

Wilson, D. A.: Doppler Radar Studies of Boundary Layer, Wind Profiles and Turbulence in Snow Conditions, Proc. 14th Conference on Radar Meteorology, Tucson, USA, p. 191–196, 1970.

Xia, Q., Lin, C. L., Calhoun, R., and Newsom, R. K.: Retrieval of Urban Boundary Layer Structures from Doppler Lidar Data. Part I: Accuracy Assessment, J. Atmos. Sci., 65, 3–20, https://doi.org/10.1175/2007JAS2328.1, 2008.

Xu, Q. and Gong, J.: Background error covariance functions for Doppler radial-wind analysis, Q. J. Roy. Meteorol. Soc., 129, 1703–1720, https://doi.org/10.1256/qj.02.129, 2002.

Zhan, L., Letizia, S., and Iungo, G. V.: LiDAR measurements for an onshore wind farm: Wake variability for different incoming wind speeds and atmospheric stability regimes, Wind Energy, pp. 1–27, https://doi.org/10.1002/we.2430, 2019.

Zhan, L., Letizia, S., and Iungo, G. V.: Wind LiDAR Measurements of Wind Turbine Wakes Evolving over Flat and Complex Terrains: Ensemble Statistics of the Velocity Field, Journal of Physics: Conference Series, 1452, https://doi.org/10.1088/1742-6596/1452/1/012077, 2020.

Zieba, A. and Ramza, P.: Standard deviation of the mean of autocorrelated observations estimated with the use of the autocorrelation function estimated from the data, Metrol. Meas. Syst., XVIII, 529–542, https://doi.org/10.1017/cbo9780511921247.018, 2011.