# Machine-learning identification of the variability of mean velocity and turbulence intensity for wakes generated by onshore wind turbines: Cluster analysis of wind LiDAR measurements

Cite as: J. Renewable Sustainable Energy **14**, 023307 (2022); https://doi.org/10.1063/5.0070094
Submitted: 03 September 2021 • Accepted: 05 April 2022 • Accepted Manuscript Online: 05 April 2022 • Published Online: 26 April 2022

G. V. Iungo, R. Maulik, S. A. Renganathan, et al.

Journal of Renewable and Sustainable Energy



AIP Publishing

Journal of Renewable and Sustainable Energy ARTICLE scitation.org/journal/rse

# Machine-learning identification of the variability of mean velocity and turbulence intensity for wakes generated by onshore wind turbines: Cluster analysis of wind LiDAR measurements

Cite as: J. Renewable Sustainable Energy **14**, 023307 (2022); doi: 10.1063/5.0070094
Submitted: 3 September 2021 · Accepted: 5 April 2022 ·
Published Online: 26 April 2022

View Online Export Citation CrossMark

G. V. Iungo,[1,a),b)] R. Maulik,[2] S. A. Renganathan,[3] and S. Letizia[1,b)]

## AFFILIATIONS

[1]Wind Fluids and Experiments (WindFluX) Laboratory, Mechanical Engineering Department, The University of Texas at Dallas, Richardson, Texas 75080, USA
[2]Argonne National Laboratory, Lemont, Illinois 60439, USA
[3]The University of Utah, Salt Lake City, Utah 84112, USA

**Note:** This article is part of the special issue Preparatory Work for the American Wake Experiment (AWAKEN).
[a)]**Author to whom correspondence should be addressed:** valerio.iungo@utdallas.edu
[b)]https://labs.utdallas.edu/windflux/

## ABSTRACT

Light detection and ranging (LiDAR) measurements of isolated wakes generated by wind turbines installed at an onshore wind farm are leveraged to characterize the variability of the wake mean velocity and turbulence intensity during typical operations, which encompass a breadth of atmospheric stability regimes and rotor thrust coefficients. The LiDAR measurements are clustered through the k-means algorithm, which enables identifying the most representative realizations of wind turbine wakes while avoiding the imposition of thresholds for the various wind and turbine parameters. Considering the large number of LiDAR samples collected to probe the wake velocity field, the dimensionality of the experimental dataset is reduced by projecting the LiDAR data on an intelligently truncated basis obtained with the proper orthogonal decomposition (POD). The coefficients of only five physics-informed POD modes are then injected in the k-means algorithm for clustering the LiDAR dataset. The analysis of the clustered LiDAR data and the associated supervisory control and data acquisition and meteorological data enables the study of the variability of the wake velocity deficit, wake extent, and wake-added turbulence intensity for different thrust coefficients of the turbine rotor and regimes of atmospheric stability. Furthermore, the cluster analysis of the LiDAR data allows for the identification of systematic off-design operations with a certain yaw misalignment of the turbine rotor with the mean wind direction.



## I. INTRODUCTION

Power generation through a wind turbine is based on the extraction of kinetic energy from the incoming atmospheric wind, which, in turn, leads to a reduced mean wind speed and enhanced turbulence intensity past the turbine rotor, namely, the generation of a wind turbine wake.[1,2]

For wind power plants, the proximity of wind turbines can lead to wake interactions for selected wind directions, namely, the wakes generated by upstream turbines can impact downstream turbine rotors, thus reducing power capture and enhancing fatigue loads of the downstream turbines. Previous field studies of onshore wind farms showed percentage power losses from 20% up to 40% of the potential power for turbines at the second row under convective conditions.[3] For stable conditions, the percentage power loss increased to 40%–60%. Moving further downstream, power losses at the third row were enhanced, especially under stable conditions, for which power losses up to 80% were observed. The cumulative energy loss of that farm due to wake interactions was estimated to be about 2.4% of the total annual energy production (AEP) when operating under convective conditions, while 4% for stable atmospheric regimes. In contrast, for offshore wind farms, the typical lower turbulence intensity of the incoming wind leads to larger downstream extent of the wind turbine

wakes,[4] and more significant power losses, such as between 10% and 20% of the AEP, as estimated for the Horns Rev wind farm,[5] or up to 28% of the nominal capacity for the Lillgrund wind farm.[6]

The near-wake characteristics, e.g., velocity deficit, wake width, and wake-added turbulence intensity, are mainly affected by the incoming wind speed, shear, and the aerodynamic characteristics of the turbine rotor, which can be encompassed by the rotor thrust coefficient. In contrast, in the far wake, i.e., downstream of the location where the maximum velocity deficit is achieved,[7] the evolution of the wake velocity field is mainly dominated by the surrounding atmospheric turbulence, which determines the intensity of the turbulent momentum fluxes promoting the gradual recovery of the wind field to the incoming wind conditions.[8–10]

Reproducing the breadth in wake variability occurring during normal operations of utility-scale wind turbines for the various parameters determining wind condition (e.g., hub-height wind speed, shear, and veer), atmospheric stability regime (e.g., turbulence intensity, Richardson number, Obukhov length), and turbine setting (e.g., rotor rotational velocity, blade pitch angle, rotor yaw angle) through numerical models and laboratory experiments is very challenging.[11] Numerical models, for instance, may struggle to generate realistic atmospheric wind conditions, which has recently motivated research to couple mesoscale and microscale models to reproduce large-scale wind heterogeneity and the corresponding smaller-scale wind turbulence.[12,13] On the other hand, modeling efficiently the action of the turbine blades on the incoming turbulent wind field for numerical simulations is still an active field of research delivering continuous improvements.[14–23]

Laboratory experiments are also becoming a fruitful resource to investigate wind turbine wakes thanks to new wind turbine models reproducing the aerodynamic forcing induced by turbine rotors,[24,25] and setups at the inlet of a test section to reproduce realistic wind conditions.[26] Nonetheless, the smaller Reynolds numbers and integral length scales of the incoming flow generated in wind and water tunnels are still important limitations to be considered for characterizing and modeling realistic velocity fields of wind turbine wakes.

Besides numerical simulations and laboratory experiments of wind turbine wakes, field observations of wakes generated by utility-scale wind turbines are becoming instrumental for learning in more detail the complex physical processes connected with wind power generation and, in turn, with the generation of wind turbine wakes.[27–30] Different remote sensing techniques, such as light detection and ranging (LiDAR),[31] radar,[32] unmanned aerial vehicles (UAVs),[33] and even instrumented larger airplanes, have been producing compelling observations of wind turbine wakes over large volumes including the entire downstream extent of wind turbine wakes, yet ensuring sufficient spatial and temporal resolutions.[34] The advancements in remote sensing for probing wind turbine wakes have involved not only the hardware and the technical aspects of the instrumentation, but also the design and the post-processing of the wind data collected under non-stationary and variable conditions, which are typical for the atmospheric wind field. Recently, a framework for the optimal design of field experiments with scanning instruments and retrieval of wind statistics, which is denoted as LiDAR statistical Barnes objective analysis (LiSBOA), has been proposed to maximize the experimental capabilities of the available remote sensing instrumentation and generate statistically accurate measurements of wind turbine wakes.[35,36]

In this paper, we aim at characterizing the variability of the wake mean velocity and added turbulence intensity under typical operations of utility-scale onshore wind turbines. Wind LiDAR measurements, together with meteorological data collected from a meteorological (met) tower, supervisory control and data acquisition (SCADA) data are leveraged for this study. Considering the large number of LiDAR samples collected for probing a turbine wake, the experimental data are first projected onto a suitable basis obtained with the proper orthogonal decomposition (POD), to reduce the dimensionality of the classification problem. Subsequently, the coefficients associated with the selected POD modes, which are considered sufficient to reconstruct the observed wake variability, are utilized for a k-means clustering algorithm to generate subsets of the initial LiDAR dataset by grouping observations ascribed to analogous atmospheric and operational conditions. Once the clustering of the LiDAR data is performed, effects on the subsets of the associated SCADA and meteorological data are also investigated, together with the mean velocity and turbulence intensity fields associated with the wakes of the various clusters.

The remainder of the paper is organized as follows: in Sec. II, the experimental dataset is described. Then, the dimensionality of the LiDAR data is reduced by applying POD in Sec. III. Subsequently, in Sec. IV the coefficients of the selected POD modes are analyzed through the k-means algorithm to generate clusters of the LiDAR wake measurements. The results of the cluster analysis on the ensemble statistics of the LiDAR measurements, SCADA, and meteorological data are then discussed in Sec. V. Finally, concluding remarks are reported in Sec. VI.

## II. LIDAR EXPERIMENT FOR AN ONSHORE WIND FARM

A LiDAR experiment was carried out at a wind farm in North Texas (the location and name of the wind farm are not disclosed per NDA agreement) comprising 39 Siemens 2.3-MW wind turbines with rotor diameter, $d$, of 108 m, a hub height of 80 m, cut-in wind speed of 3 m/s, rated wind speed of 11.5 m/s, and cutout wind speed of 25 m/s. The topography map of the site is retrieved from the U.S. Geological Survey[37] with a spatial resolution of 100 m [Fig. 1(a)]. The standard deviation of the terrain is only 16 m, which allows considering this site as flat terrain. For the retrieval of the LiDAR data, the hub height of each turbine is corrected by taking the local altitude at the turbine locations into account.

The measurement campaign was conducted through various phases between August 2015 and March 2017 for a total of 236 days. Meteorological data were provided as 10-minute averages and standard deviation of wind speed, wind direction, temperature, humidity, and barometric pressure at heights of 36, 60, 75, and 80 m. Based on data availability and the quality control of the meteorological data, wind speed and direction were available for the duration of the experiment only for the 60-m height, while temperature data at 75-m height. The atmospheric stability regime is characterized through the Bulk Richardson number[38]

$$Ri_B(\bar{z}) = \frac{g\,\Delta T/\Delta z}{T(z_w)\,U^2(z_w)}\bar{z}^2, \tag{1}$$

where $g$ is the gravitational acceleration, $z_w = 60$ m is the met-tower height where the wind speed, $U$, is measured, $\Delta T$ is the temperature variation over $\Delta z = z_2 - z_1$ ($z_2 = 75$ m and $z_1 = 3$ m), and $\bar{z} = \sqrt{z_1 z_2} = 15$ m. Wind data at 3-m height were collected from a CSAT3 3D

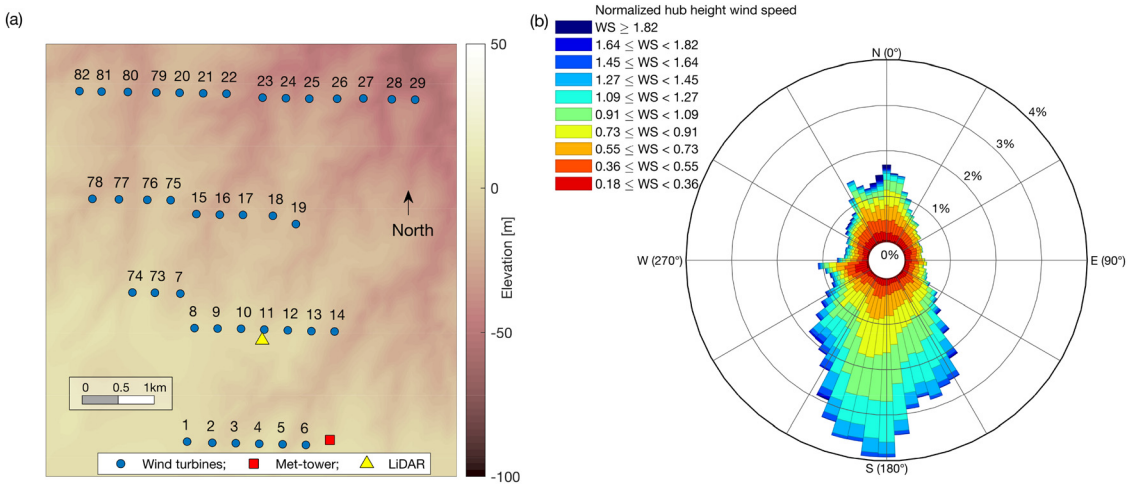


**FIG. 1.** Test site: (a) layout of the wind farm, where the size of the blue markers represents down-scaled rotor diameter; (b) wind rose of the hub-height wind measured by the met-tower for the entire duration of the experiment and reported as a ratio of the turbine rated wind speed, which is 11.5 m/s.

sonic anemometer manufactured by Campbell Scientific Inc. installed on a mast located in the proximity of the LiDAR.

SCADA data were provided for each turbine as 10-minute averages and standard deviation of wind speed, power output, rotor rotational velocity, and pitch angle. For more details of this dataset and the used quality control process, see El-Asha, Zhan, and Iungo[3] and Zhan, Letizia, and Iungo.[9]

The scanning pulsed Doppler wind LiDAR deployed for this experiment, whose location is indicated with a yellow triangle in Fig. 1(a), is a Windcube 200S manufactured by Leosphere, which emits a laser beam into the atmosphere and measures the radial wind speed, i.e., the velocity component parallel to the laser beam, from the Doppler frequency shift of the back-scattered LiDAR signal. According to the wind farm layout and the prevalence of southerly wind directions [Fig. 1(b)], for wind directions within the sector between 145° and 235°, the wakes produced by the turbines from 1 to 6 evolve roughly toward the LiDAR location, which is a favorable condition for the LiDAR to measure with close approximation the streamwise velocity through plan-position indicator (PPI) scans. Furthermore, according to the layout of Fig. 1(a), for the considered wind directions, these wind turbines are not affected by upstream wakes.

The LiDAR measurements were typically performed using a range gate of 50 m, an elevation angle of $\phi = 3°$, an azimuthal range of 20°, and a rotation speed of the scanning head of 2°/s, leading to a typical scanning time for a single PPI of 10 s. After rejecting LiDAR data with a carrier-to-noise ratio (CNR) lower than −25 dB, a proxy for the streamwise velocity is obtained through the streamwise equivalent velocity,

$$U_{eq} = V_r/[\cos\phi\ \cos(\theta - \theta_w)], \tag{2}$$

where $V_r$ is the LiDAR radial velocity, $\theta$ is the azimuthal angle of the LiDAR laser beam, and $\theta_w$ is the wind direction. Considering that the PPI LiDAR data are collected at different heights, the equivalent velocity is made non-dimensional with the incoming boundary layer profile, which is estimated from the PPI scans through the value of the 70th percentile of the distribution of the streamwise equivalent velocity for each height probed by the LiDAR.[9]

The reference frame used has the $x$-direction aligned with the wake direction, which is estimated with the linear fitting of the wake centers at various downstream locations, the $y$-direction in the horizontal transverse direction, and $z$-direction vertically and positive moving upward. The transverse position of the wake center is defined as the location of the minimum velocity obtained by fitting the velocity data at a specific downstream distance through a Gaussian function. More details of the LiDAR system, the field campaign, and data post-processing are available in Zhan, Letizia, and Iungo.[9]

Total number $N = 6654$ quality-controlled PPI LiDAR scans of isolated wind turbine wakes have been processed to provide the non-dimensional wake velocity fields used for this study.[9,10] To estimate the normalized velocity field relevant for wake modeling, a generic LiDAR sample collected at the location $(x, y, z)$ is ascribed to the coordinates $(x, r)$, where the coordinate, $r$, is defined as

$$r = sign(y)\sqrt{y^2 + z^2}, \tag{3}$$

noting that $r$ has a sign to identify the two sides from the $x$-axis of the wake. For the wake analysis, only LiDAR samples with $|z/d| < 0.25$ are considered to ensure that they are representative of the wake velocity field at hub height.

Considering that the $x$-direction of the reference frame is aligned with the wake direction, it entails that the Cartesian coordinates of a given LiDAR sample can vary for the various scans due to changes in wake direction. Therefore, for each scan, the LiDAR data need to be interpolated over a common Cartesian grid to calculate ensemble flow statistics for each data cluster. Specifically, this Cartesian grid is defined over the horizontal plane at hub height with dimensions $1 \le x/d \le 7$ and $-1 \le r/d \le 1$. LiDAR data collected for $x/d > 7d$ are not considered because acquired at heights $z/d < -0.25$ and, thus, they might reduce accuracy in the calculations of the velocity statistics at the hub height. LiDAR data collected at $x/d < 1$ are rejected

because they are typically affected by spurious contamination of the velocity signals due to the LiDAR laser beam hitting turbine components. The wake velocity data have a spatial resolution of $0.1d$ and $0.05d$ in the $x$ and $r$ directions, respectively, generating a data matrix of $[p \times q] = [61 \times 41]$ ($pq = 2501$) for a single LiDAR snapshot.

For grid points where the LiDAR data are not available, which can be due to the quality control process of the LiDAR data or a significant misalignment between the wake direction and the direction connecting the LiDAR and turbine locations, the velocity fields are interpolated through the inpaint-nans function available in Matlab.[39] The ensemble-averaged velocity fields calculated over the entire dataset are reported in Fig. 2 for both raw and interpolated data. The main data distortion due to the velocity interpolation occurs at the downstream corners of the spatial domain, which are the areas where the probability of missing LiDAR samples is higher. The interpolated velocity fields are only used for the POD, which does not allow for not-a-number (NaN) values over the spatial domain for the calculation of the eigenproblem of the velocity covariance matrix. In contrast, statistics of the wake velocity field will be calculated with the original non-interpolated velocity fields by rejecting NaN values.

## III. PROPER ORTHOGONAL DECOMPOSITION OF THE LIDAR DATASET

The analysis of a large dataset encompassing correlated variables aiming to detect trends, data variability, and features can be very challenging and computationally demanding. Specifically, considering a cluster analysis with a total number of $K$ clusters ($K \approx 10$ for this work), $I$ iterations for the k-means algorithm ($I \approx 10$ K), and $L$ repetitions ($L \approx 100$), the total number of integrals to be computed is $L \times I \times K \times N \sim \mathcal{O}(10^9)$.[40] Therefore, rather than analyzing the LiDAR dataset in its entirety, it can be more convenient to represent the data on a lower-dimensional subspace with a suitable lower rank to enable a simplified analysis, which typically entails lower computational costs and clearer interpretation of the results.[41–43] Specifically, POD[44–46] allows for the generation of an orthonormal basis, which is optimal for the reconstruction of the data variability. The most computationally expensive task of POD is the calculation of the correlation matrix of the LiDAR measurements [proportional to $N \times (N+1)/2$]. Therefore, the computational saving is about two orders of magnitude $(N+1)/(2L \times I \times K) = \mathcal{O}(10^{-2})$.

POD is computed for the non-dimensional interpolated LiDAR data with the method of snapshots.[47] A snapshot of the wake velocity field measured with the LiDAR, $u$, can be represented through a linear combination of deterministic functions, which are referred to as POD modes $\phi_j$

$$u(x, r, t) = \sum_{j=1}^{pq-1} \phi_j(x, r)\, a_j(t), \tag{4}$$

where $t$ is time. The POD modes, $\phi_j$, represent the typical spatial variability of the velocity field in a statistical sense, while the parameters $a_j(t)$ are coefficients representing the amplitude of each POD mode as a function of time. POD provides a modal decomposition that is completely a-posteriori and data-dependent, which does not neglect the non-linearities of the original dynamical system, even being a linear procedure. Furthermore, the POD basis is orthonormal and optimal in variance, i.e., among all linear decomposition techniques, it provides the most efficient detection, in a certain least squares optimal sense, of the dominant components.

The LiDAR dataset, $\mathbf{U}$, with dimensions $[pq \times N]$ can be approximated by computing the first $r$ most energetic principal components through the singular value decomposition (SVD)

$$\mathbf{U} \approx \Phi \times \Sigma \times V^{\top}, \tag{5}$$

where $\Phi$ ($[pq \times r]$) and $V$ ($[N \times r]$) are orthonormal matrices and $\Sigma$ is a diagonal matrix ($[r \times r]$) with the first $r$ singular values of $\mathbf{U}$ in descending order as diagonal entries.[48] Each diagonal element of $\Sigma$, $\sigma_j$, represents the energy contribution of the POD mode $\phi_j$ to the covariance matrix of the velocity snapshots. The POD modes are obtained as columns of the matrix $\Phi$, while the principal components, $a_j$, associated with each POD mode are obtained by projecting the snapshot dataset onto the POD basis

$$\mathbf{A} = \mathbf{U}^{\top}\Phi, \tag{6}$$

where the principal components, $a_j$, are the columns of $\mathbf{A}$, whose size is $[N \times r]$.

POD is applied to the non-dimensional interpolated LiDAR wake measurements collected over the horizontal plane at hub height. It is now crucial to select the smallest number of POD modes enabling an efficient reconstruction of the wake variability probed through the LiDAR measurements. The selection of POD modes can be done based on the energetic contribution of the various POD modes. In other words, by leveraging the energy optimality of the POD basis, the latter is truncated to reconstruct a certain percentage of the total energy of the velocity covariance matrix. An alternative to this energetic approach has been used for the present work, which consists of visually inspecting the most energetic POD modes and selecting only POD modes whose spatial morphology indicates a clear physical feature.[43] It is noteworthy that while the visual inspection of the POD modes can increase the effectiveness for the POD-based reconstruction

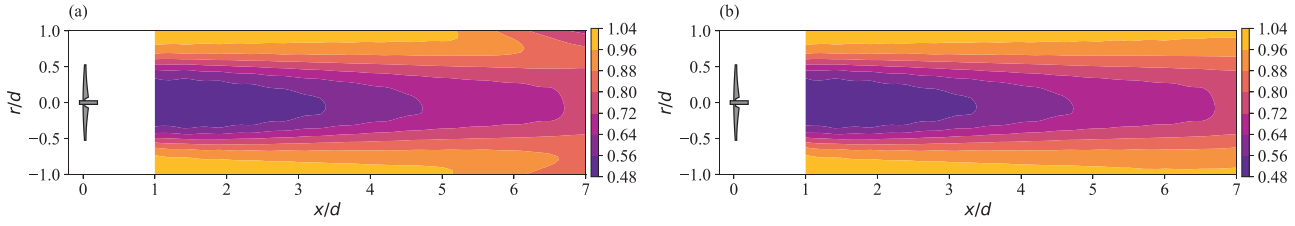


**FIG. 2.** Ensemble mean of the non-dimensional wake velocity fields calculated over the entire dataset: (a) NaN values interpolated with the function inpaint-nans; (b) interpolation-free data.

of the wake velocity field, on the other hand, it makes the procedure less automated because requiring direct inputs from the user.

The first 12 most energetic POD modes are reported in Fig. 3. POD mode 0 resembles the ensemble average of the interpolated wake velocity fields, which is shown in Fig. 2(a). Therefore, this POD mode has an evident physical contribution to the wake morphology. In contrast, POD modes 1 and 2 represent the corrections performed through the interpolation of the LiDAR data with the Matlab function inpaint-nans,[39] and, thus, they are ignored being a numerical artifact rather than a physical feature for the POD basis. POD modes 3 and 4 seem to indicate a modulation in the transverse direction over the shape provided through the ensemble mean (POD mode 0). In other words, POD modes 3 and 4 can represent contractions or expansions of the wake in the transverse direction,[49] non-symmetric wake conditions that were already observed in previous works for the near wake under stable atmospheric conditions,[9,10] wake meandering,[50] or lateral deflections due to the Coriolis effect.[51] Similarly, POD modes 5 and 11 seem to indicate contractions or extensions of the wake in the streamwise direction. The remaining POD modes shown in Fig. 3 seem to indicate similar wake distortions, yet with slightly larger wavelengths, which can be considered as sub-harmonics of the above-mentioned wake modulations. Based on this qualitative analysis of the POD modes, which we understand is speculative rather than based on quantitative characteristics, the truncated POD basis selected for this work includes only POD modes 0, 3, 4, 5, and 11, which allows reconstructing 44.8% of the overall energy of the velocity covariance matrix.

The approximation of the LiDAR data with the POD projection typically leads to a smoothed wake morphology throughout the spatial domain, which seems an affordable drawback of the POD data compression, considering the reduced computational costs of the following cluster analysis achieved by reducing the dimensionality of each snapshot from $pq = 2,501$ down to 5.

## IV. k-MEANS CLUSTERING OF THE WIND LIDAR DATA

Clustering, in general, refers to a very broad set of techniques for finding subgroups from a dataset, which are referred to as clusters.[52] The aim of clustering a dataset is to partition it into distinct sub-groups where samples sharing similar features belong to the same cluster and are segregated from those characterized by different features.

For this work, the *k*-means algorithm is used, which is a simple approach for partitioning a dataset into $K$ distinct, non-overlapping clusters.[53] To perform *k*-means clustering, the desired number of clusters, $K$, must be provided as input. The standard method used to perform *k*-means clustering is an iterative algorithm. First, a random integer (from 1 to $K$) is assigned to each sample as an initial cluster assignment. Next, for each of the $K$ clusters, the centroid of the cluster is computed and each observation is assigned to the cluster that it is closest to. These two steps are repeated until convergence in the cluster association is achieved.[48]

After the generation of clusters, a silhouette analysis is performed for each cluster to reject samples considered as outliers for their respective cluster.[54] The silhouette analysis is performed by quantifying for each sample its distance from the respective cluster centroid through the $L_2$-norm and the distances from the centroids of the remaining clusters. The silhouette coefficient has a range between

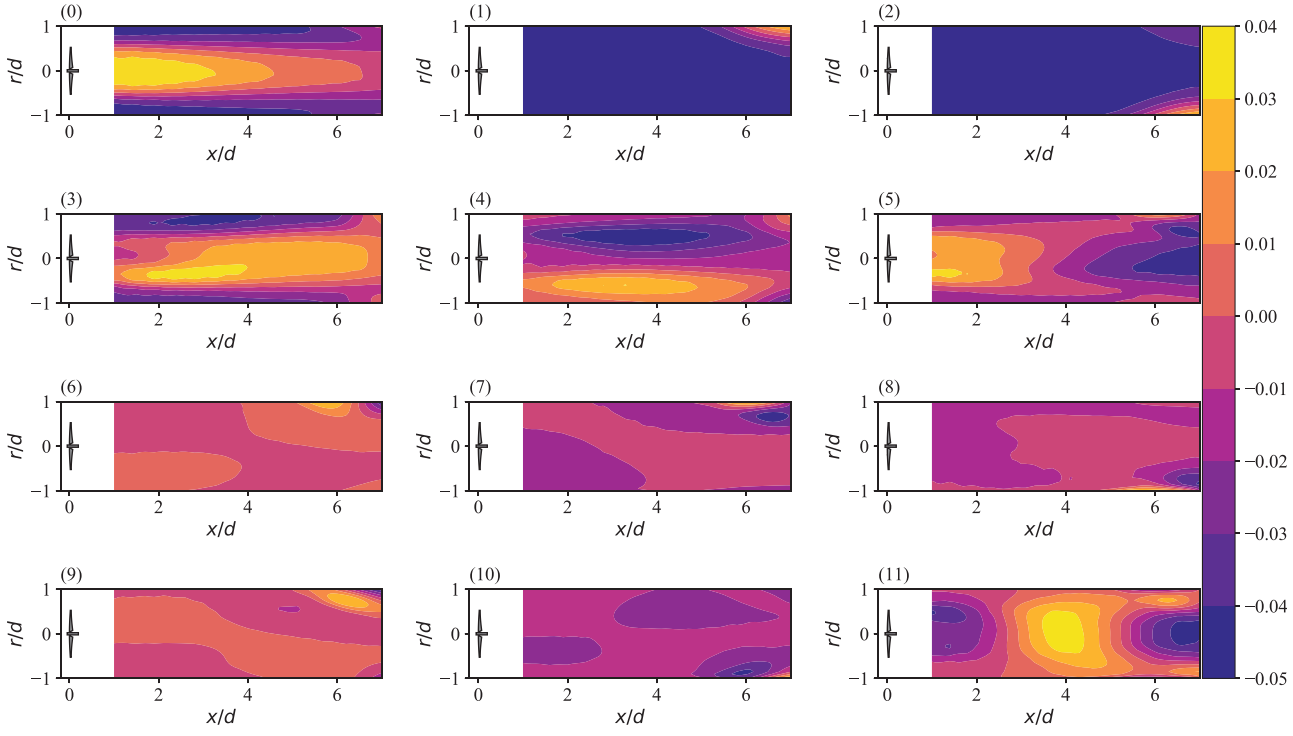


**FIG. 3.** First 12 most energetic POD modes obtained from the LiDAR dataset.

−1 and 1, where 1 corresponds to the respective cluster centroid, 0 to the boundary of the cluster associated with the sample, and a negative number indicates that the sample might be an outlier for the considered cluster. In this work, all the samples with a non-positive silhouette coefficient are rejected for further analyses, which correspond to 7% of the source dataset.

For our study, the only inputs provided for the cluster analysis are the time-series of the coefficients associated with the five POD modes selected to approximate the LiDAR dataset. It is noteworthy that no input is related neither to the wind turbine operative conditions nor to the wind/atmospheric conditions, as for previous cluster analyses of wind turbine wake measurements.[9,10] The *k*-means outputs are the cluster centers, namely, the representative realization for each cluster, and labeling for each LiDAR snapshot to its respective cluster.

One of the challenges with using the *k*-means clustering algorithm is the choice of the number of clusters, *K*. This decision process is facilitated through the evaluation of the inertia parameter, which is one of the outputs of the *k*-means algorithm indicating the statistical relevance of each cluster within the dataset.[93] It should be considered that increasing the number of clusters, *K*, might lead to the partitioning of clusters with a larger data population, rather than identifying different data features. Therefore, for this work, we have applied a hierarchical clustering approach, where the data are classified according to a cluster tree, denoted as dendrogram.[55] The dataset and the generated clusters (denoted as nodes) are partitioned into more successor sub-groups. Finally, all the nodes and sub-groups are nested and organized with a tree-like structure to provide a more physical and meaningful classification of the data. For the sake of efficiency, the number of the groups generated from a single node should be limited to clusters representing distinct physical features, i.e., wake characteristics in terms of velocity deficit and recovery rate.

After a preliminary cluster analysis, we decided to perform a first partitioning of the LiDAR dataset into three clusters, denoted as *Ca*, *Cb*, and *Cc*. The centroids of these clusters are reported in Fig. 4, which are obtained by adding the centroids of each principal component multiplied by their respective POD mode. In other words, these cluster centroids are a proxy for the ensemble average of the LiDAR velocity fields belonging to the same cluster.

With this initial clustering, the dataset has been partitioned within groups with similar populations, as reported in the row "Occurrence" of Table I. Figure 4 shows that the main differences among the centroids of the three clusters are related to the wake velocity deficit and the downstream extent of the wakes.

These three clusters (*Ca*, *Cb*, *Cc*) are now considered as nodes of the dendrogram, and, thus, each cluster is re-processed through the *k*-means algorithm to generate further sub-groups of the LiDAR dataset. The cluster analysis of *Ca* has led to the generation of subgroups (not shown here for the sake of brevity) with very

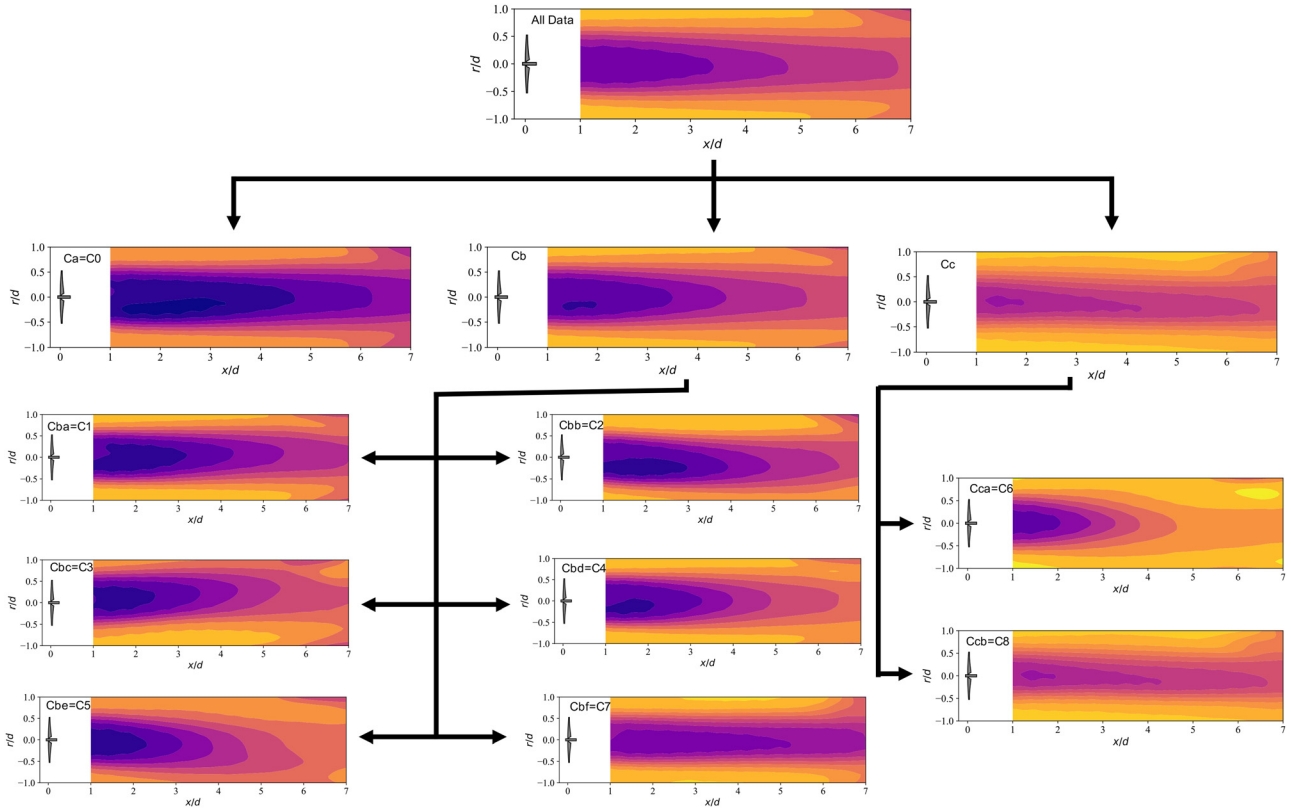


**FIG. 4.** Dendrogram of the LiDAR dataset reporting the cluster centers. The iso-curves represent values of $U/U_\infty$ from 0.4 up to 1.1 with steps of 0.1.

**TABLE I.** Cluster parameters.

| Cluster label | C0 | ··· | ··· | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Initial label | *Ca* | *Cb* | *Cc* | *Cba* | *Cbb* | *Cbc* | *Cbd* | *Cbe* | *Cca* | *Cbf* | *Ccb* |
| Occurrence | 28.7% | 40.4% | 30.8% | 9.1% | 5.3% | 4.8% | 6.2% | 3.3% | 4.0% | 8.2% | 23.4% |
| Inertia | 35.2 | 19.1 | 22.3 | 8.3 | 10.2 | 17.1 | 9.1 | 7.7 | 15.6 | 12.8 | 19.9 |
| $x_{tr}/d$ | 2.9 | ··· | ··· | 2.2 | 2.3 | 1.5 | 1.9 | 1.5 | 1.4 | 2.3 | 1.9 |
| $A_u$ | 1.00 | ··· | ··· | 1.38 | 1.32 | 1.55 | 2.17 | 2.20 | 3.00 | 0.59 | 0.49 |
| $N_u$ | 0.46 | ··· | ··· | 0.78 | 0.83 | 1.10 | 1.29 | 1.45 | 2.26 | 0.23 | 0.34 |
| $v_t/U_\infty/d$ | 0.0034 | ··· | ··· | 0.0073 | 0.0105 | 0.0223 | 0.0152 | 0.0305 | 0.1000 | 0.0010 | 0.0021 |
| $C_P$ | 0.35 | ··· | ··· | 0.40 | 0.40 | 0.39 | 0.41 | 0.46 | 0.43 | 0.30 | 0.26 |
| $C_T^{\mathrm{SCADA}}$ | 0.40 | ··· | ··· | 0.46 | 0.46 | 0.45 | 0.47 | 0.55 | 0.51 | 0.33 | 0.28 |
| $C_T^{AD}$ | 0.95 | ··· | ··· | 0.90 | 0.89 | 1.01 | 0.87 | 1.14 | 0.81 | 0.70 | 0.57 |
| $k^*$ | 0.017 | ··· | ··· | 0.029 | 0.017 | 0.050 | 0.023 | 0.071 | 0.050 | 0.012 | −0.004 |
| $\sigma_{U_\infty}/U_\infty$ | 0.04 | ··· | ··· | 0.066 | 0.063 | 0.085 | 0.078 | 0.082 | 0.088 | 0.039 | 0.053 |

similar flow characteristics. Therefore, cluster *Ca* is considered a definitive cluster for our study, which is ultimately labeled as *C*0. In contrast, the sub-cluster analysis of node *Cb* leads to the identification of subgroups with a gradually varying downstream extent of the wake. In Fig. 4, these sub-groups are indicated from *Cba* to *Cbf*. A total number of six sub-groups is selected after a preliminary analysis to cover the variability of the wake velocity field through the cluster centers while limiting the number of clusters generated.

For the cluster analysis of the node *Cc*, two clusters are deemed sufficient to identify the main wake topologies encompassed within this sub-dataset. Specifically, the sub-group *Cca* is characterized by a strong near-wake velocity deficit and fast recovery, while the sub-group *Ccb* preserves wake features similar to those of the respective node *Cc*.

In summary, through the *k*-means algorithm and the dendrogram approach, nine clusters are generated from the initial LiDAR dataset, which are numbered from *C*0 up to *C*8. The ensemble statistics in terms of average and standard deviation for the entire dataset and the various clusters are reported in Fig. 5. As mentioned above, the ensemble mean fields resemble the cluster centroids already reported in Fig. 4, yet avoiding data distortion introduced by the interpolation of the LiDAR data required for the application of the POD. The ensemble statistics highlight the broad variability of the wake velocity field for different settings of the turbines and meteorological conditions, both for the mean and standard deviation. Furthermore, it is noteworthy that clusters *C*2 and *C*3 show evident skewing of the wake from the *x*-axis representing the wake direction. This particular wake morphology, which can be ascribed to a certain yaw misalignment of the rotor disk from the mean wind direction,[56] will be analyzed more in detail in Sec. V.

## V. ANALYSIS OF THE CLUSTERED DATASETS

The statistical analysis presented in this section is performed with the original non-dimensional LiDAR data partitioned according to the nine clusters obtained in Sec. IV, thus, avoiding any spurious effect on the velocity statistics that might be introduced in case the POD-projected LiDAR data were used.

### A. Mean wake velocity field for the various clusters

For characterizing the mean wake velocity obtained for the various clusters, the maximum velocity deficit, $\Delta U_{min}/U_\infty = (U_\infty - U_{min})/U_\infty$, ($U_{min}$ is the minimum streamwise velocity at a given downstream distance), as a function of the downstream position is reported in Fig. 6. This plot emphasizes even more clearly the variability in velocity deficit and wake persistence along the downstream direction that can be observed during typical operations of a wind turbine. For the clusters from *C*0 up to *C*6, the velocity deficit in the very near wake is comparable, which suggests that the thrust coefficient of the turbine might be very similar for these clusters. Ranging from *C*0 to *C*6, we generally observe a faster recovery of the velocity deficit. The velocity deficit at $x/d = 1$ is lower for cluster *C*7 and even more for *C*8. This wake feature may indicate that clusters *C*7 and *C*8 belong to turbine operations with an active pitch control of the turbine blades and, thus, reduced rotor thrust coefficient. For *C*7 and *C*8, the recovery of the velocity deficit is significantly slower than for the remaining clusters.

Starting from the most upstream location, i.e., $x/d = 1$, the velocity deficit generally increases in the downstream direction, which indicates that the turbine forcing, mainly through the generation of a streamwise pressure gradient, is still acting on the incoming flow.[57] The downstream location where the maximum velocity deficit occurs, $x_{tr}$, can be associated with the transition between the near wake and the far wake.[7,58] This analysis shows that the near-to-far wake transition occurs for downstream distances in the range between 1.4*d* and 2.9*d*, which is in agreement with previous studies,[59] and generally moves upstream switching from cluster *C*0 to *C*6, while it seems roughly invariant for *C*7 and *C*8 (Table I).

It is noteworthy how the design of improved LiDAR scans coupled with the data-driven clustering of the LiDAR data enables more accurate characterization of the near-wake velocity field, which is associated with the pressure gradients induced by the turbine rotor. The maximum velocity deficit is typically achieved at a downstream distance smaller than 3*d*,[59] while further downstream turbulent mixing governs far-wake recovery.[57]

To provide a more quantitative analysis of the mean wake velocity fields associated with the various clusters, the maximum velocity

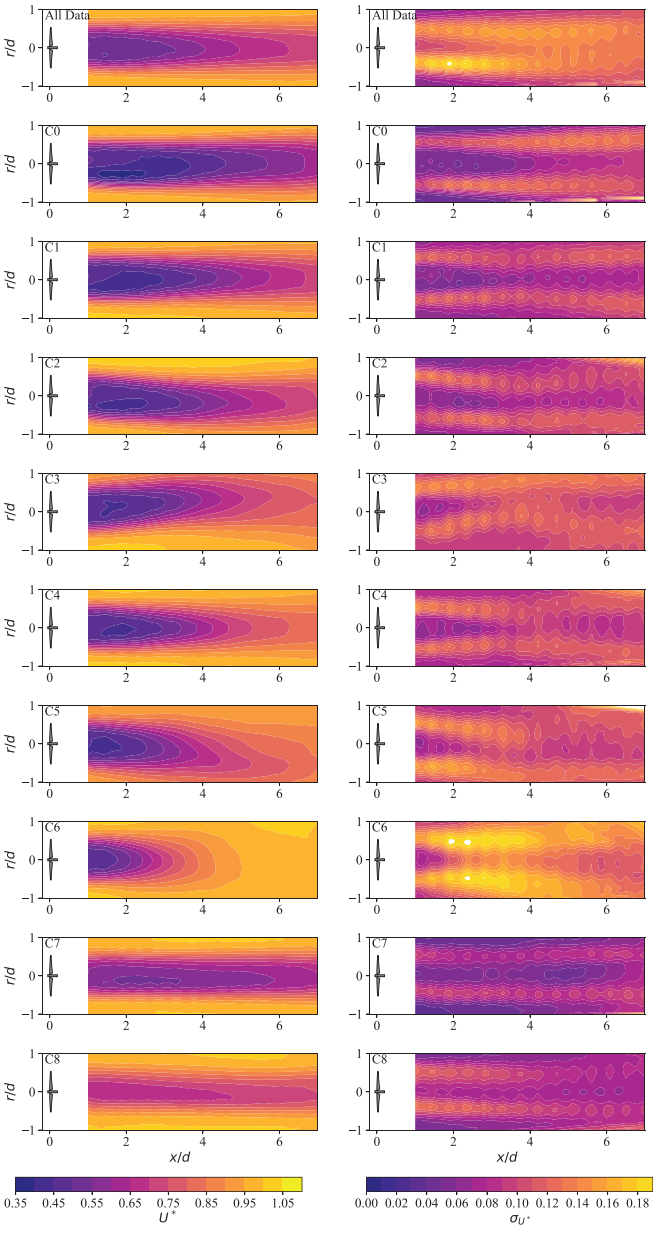


**FIG. 5.** Ensemble statistics of the LiDAR wake measurements for the various clusters: (left column) mean non-dimensional velocity fields, $U/U_\infty$; (right column) standard deviation of the non-dimensional velocity fields, $\sigma_U/U_\infty$.

deficit, $\Delta U_{min}/U_\infty$, is fitted through the following power law for $x/d \geq 3$:[8,9]

$$\frac{\Delta U_{min}}{U_\infty} = 1 - A_u \left(\frac{x}{d}\right)^{-N_u}, \tag{7}$$

where $A_u$ and $N_u$ represent the velocity deficit at $x/d = 1$ and the wake-recovery rate, respectively. As reported in Table I, the parameter $A_u$ generally increases spanning from the cluster $C0$ up to $C6$ between

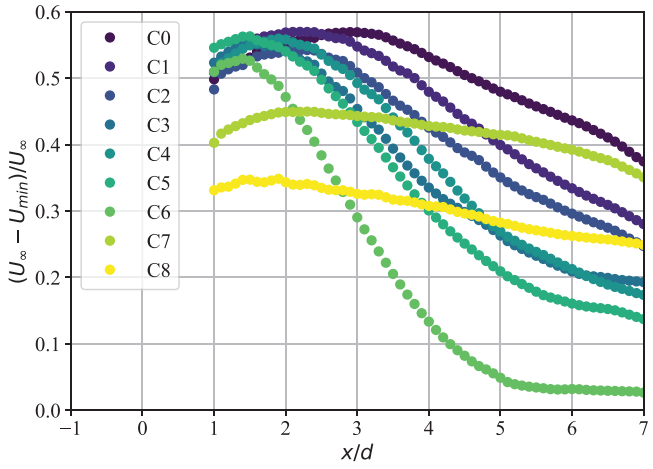


**FIG. 6.** Maximum velocity deficit, $\Delta U_{min}$, as a function of the streamwise coordinate for the various clusters.

values 1 and 3, respectively. The fitted values of $A_u$ are significantly larger than the actual maximum velocity deficit at $x/d = 1$ because of a different variability of the mean velocity field in the near wake, i.e., for $x \approx< x_{tr}$ (see Fig. 6). Similarly, the parameter $N_u$ increases from 0.46 for cluster $C0$ up to 2.26 for $C6$. Therefore, the variability of $N_u$ corroborates a faster wake recovery spanning from $C0$ up to $C6$.

Completely different wake features are quantified for clusters $C7$ and $C8$, for which the reduced near-wake velocity deficit leads to $A_u$ equal to 0.59 and 0.49, respectively, and very low wake recovery rates than for the previous clusters, i.e., 0.23 and 0.34, respectively.

As mentioned above, the velocity deficit in the near-wake is strictly connected with the rotor thrust coefficient, $C_T$.[10,60] Estimates of $C_T$ from the mean wake velocity field can be generated by considering a 1D stream tube with the bases centered along the $x$-direction and located upstream and downstream of the rotor disk. By coupling mass conservation and the momentum equation in the streamwise direction, it is obtained[61,62]

$$C_T^{AD} = \frac{8}{\pi} \int_0^{2\pi} \int_0^{\infty} \frac{U}{U_\infty} \left(1 - \frac{U}{U_\infty}\right) \frac{r^*}{d}\, d\theta\, d\left(\frac{r^*}{d}\right), \tag{8}$$

where $(r^*, \theta)$ are the cylindrical coordinates associated with the stream tube. Equation (8) is solved numerically for the mean velocity fields $U/U_\infty$ of each cluster at the various downstream locations by injecting the velocity field from one side of the wake with positive $r$ for the range $0 \leq \theta < \pi$, and the velocity field on the other side of the wake with negative $r$ for the range $\pi \leq \theta < 2\pi$. The estimates of the thrust coefficient through the actuator disk theory, $C_T^{AD}$, are obtained for each cluster at the location of the maximum velocity deficit, $x_{tr}$ (see Table I). These results confirm that the thrust coefficient for clusters $C0$–$C6$ is substantially larger (between 0.81 and 1.14) than for $C7$ and $C8$ (0.7 and 0.57, respectively).

It is noteworthy that for the clusters $C0$–$C6$, the variability of $C_T^{AD}$ does not reflect the variability observed for the parameter $A_u$. Indeed, the quantification of $A_u$ is affected not only by the rotor thrust in the near wake but also by the location of the transition from the near and far wake, $x_{tr}$. Therefore, we consider $C_T^{AD}$ a more reliable

parameter for the characterization of the rotor thrust and the near-wake velocity field.

The recovery rate of the ensemble mean velocity field for each cluster is further characterized by calculating the optimal turbulent eddy-viscosity, $\nu_T$, for an axisymmetric Reynolds-averaged Navier–Stokes (RANS) model[60] (the RANS code is publicly available[63]) Specifically, for each cluster, the mean velocity measured at $x/d = 1$ is provided as the inlet condition for the RANS model, while a constant $\nu_T$ is optimized by minimizing the $L_2$-norm of the difference between the LiDAR mean velocity field and the velocity field predicted through the RANS model. For the optimization, the sequential quadratic programming implemented in Matlab through the fmincon function is adopted[64] with termination tolerance on the first optimality of $10^{-6}$. The estimates for $\nu_T$ show similar trends as for the parameter $N_w$, namely, the lowest non-dimensional values are obtained for the clusters *C*7 and *C*8 (0.0010 and 0.0021, respectively), reflecting the very slow wake recovery and long downstream extent of the respective mean velocity fields. A larger $\nu_T/U_\infty/d$ is estimated for *C*0 (0.0034), which is characterized by a significant downstream extent, yet larger near-wake velocity deficit. For the remaining clusters, $\nu_T/U_\infty/d$ increases up to 0.1, reflecting a gradually shorter downstream extent and faster wake recovery.

The majority of the clusters are characterized by a quasi-symmetric wake. However, the clusters *C*2 and *C*3 show a significant transverse deflection of the wake center, reaching maximum values, $\delta_\infty$, of about $-0.19d$ and $0.33d$, respectively (Fig. 5). This wake feature can be ascribed to a misalignment of the rotor axis with the mean wind direction of a certain angle $\gamma$, similarly to a wind turbine operating under yawed conditions.[56,65,66] However, it is known that the wind turbines were supposed to operate without any intentional yaw misalignment. For a given asymptotic wake deflection, the rotor yaw angle, $\gamma$, can be estimated as[56]

$$\frac{\delta_\infty}{d} = \frac{0.3\gamma}{\cos\gamma}\left(1-\sqrt{1-C_T\cos\gamma}\right)\left[\frac{\cos\gamma\left(1+\sqrt{1-C_T}\right)}{\sqrt{2}\left(\alpha^* TI+\beta^*\left(1-\sqrt{1-C_T}\right)\right)}+\frac{1}{14.7}\sqrt{\frac{\cos\gamma}{k^{*2}C_T}}\left(2.9+1.3\sqrt{1-C_T}-C_T\right)\times\ln\left(\frac{1.6+\sqrt{C_T}}{1.6-\sqrt{C_T}}\right)\right], \tag{9}$$

where $\alpha^* = 2.32$, $\beta^* = 0.154$, and $k^*$ is the linear growth rate of the wake width. Using as the thrust coefficient the values obtained through the actuator disk theory, $C_T^{AD}$, and $k^*$ estimated through the Gaussian model[67] (both parameters are reported in Table I), these asymptotic wake deflections would correspond to a rotor yaw angle, $\gamma$, of $-3.2°$ and $5.6°$ for *C*2 and *C*3, respectively. Unfortunately, the actual yaw angle of the wind turbines is not provided from the SCADA data, and it is known that the wind turbines under investigation were supposed to operate with a zero yaw angle. Therefore, the observed skewed morphology of the wake in the lateral direction for the clusters *C*2 and *C*3 seems to be associated with a variability of the wind direction along the rotor heights or non-ideal performance of the sensors/controller of the wind turbines in setting the rotor yaw angle.

This result is interesting because the cluster analysis of the LiDAR data has enabled the identification of these operations with a systematic yaw misalignment, leading to a noticeable deflection of the wind turbine wakes, which might be important for the power efficiency of the entire wind farm. These operations under yaw misalignment have a total occurrence of about 10% of the duration of the LiDAR observations, which is not a negligible number. The detection of these off-design conditions would not have been possible by generating bins of the LiDAR dataset by imposing thresholds on the various wind and turbine parameters, as done for previous works;[9] in contrast, this result has been achieved only thanks to the data-driven approach of the k-means clustering algorithm.

### B. Effects of the clustering analysis on the meteorological and SCADA data

For the various clusters, we now analyze the statistics of SCADA and meteorological parameters, which were collected simultaneously to the LiDAR data. It is noteworthy that these parameters were not fed into the cluster analysis and, thus, their variability has not affected the generation of the data clusters. The histogram of the mean hub-height wind speed recorded through the SCADA is reported in Fig. 7 for the entire dataset and the various clusters, while the 25, 50, and 75 percentiles for the entire dataset and the various clusters are reported in Table II. Considering that the rated wind speed of the turbines under investigation is 11.5 m/s, then it becomes evident that the clusters from *C*0 up to *C*6 belong to operations in region 2 of the power curve, while the cluster *C*8 belongs to operations in region 3, i.e., between rated wind speed and cutoff wind speed, and the cluster *C*7 seems to encompass operations at the transition between region 2 and region 3, typically denoted as region 2.5.[68]

This classification of the various clusters based on the turbine operative conditions is corroborated from the respective histograms of the turbine power and blade pitch angle, Figs. 8 and 9, respectively, and related statistics reported in Table II. Power capture for the clusters *C*0–*C*6 is lower than the rated power of 2300 kW, while the blade pitch angle is generally between $-2°$ and $-1.5°$, which is the typical range for non-active pitch control of the turbine rotor. In contrast, power capture for the operations of clusters *C*7 and *C*8 belongs to operations with an active blade pitch control. Indeed, the statistics of the pitch angle for *C*7 and *C*8 are larger than for *C*0–*C*6 with median values of $1.26°$ and $2.64°$, which are associated with a median power capture of 2239 and 2304 kW, respectively.

Besides turbine settings, also atmospheric stability can significantly affect the characteristics and evolution of wind turbine wakes.[8,69–73] To investigate the role of atmospheric stability on the variability observed through the ensemble statistics of the wake velocity field for the various clusters, histograms (Figs. 10–12) and percentiles (Table II) of the Bulk Richardson number, $Ri_B$, incoming wind turbulence intensity at hub height, *TI* (defined as the ratio between the standard deviation and the mean value of the streamwise velocity), and shear exponent, $\alpha$, are also analyzed, respectively. Starting from the clusters owing operations in region 2 of the power curve (*C*0–*C*6), the Bulk Richardson number is generally positive for *C*0 (median $Ri_B$ of 0.0016), while it is typically negative for the remaining clusters *C*1–*C*6. Therefore, operations associated with *C*0 can occur under stable atmospheric conditions while *C*1–*C*6 under unstable conditions. Furthermore, the level of atmospheric instability increases ranging from *C*0 up to *C*6 (Table II).

This analysis is confirmed from the statistics of the incoming wind turbulence intensity at hub height, *TI*, which has a median value of 5.8% for *C*0 and the largest value of 15.9% for *C*6. Similar conclusions are

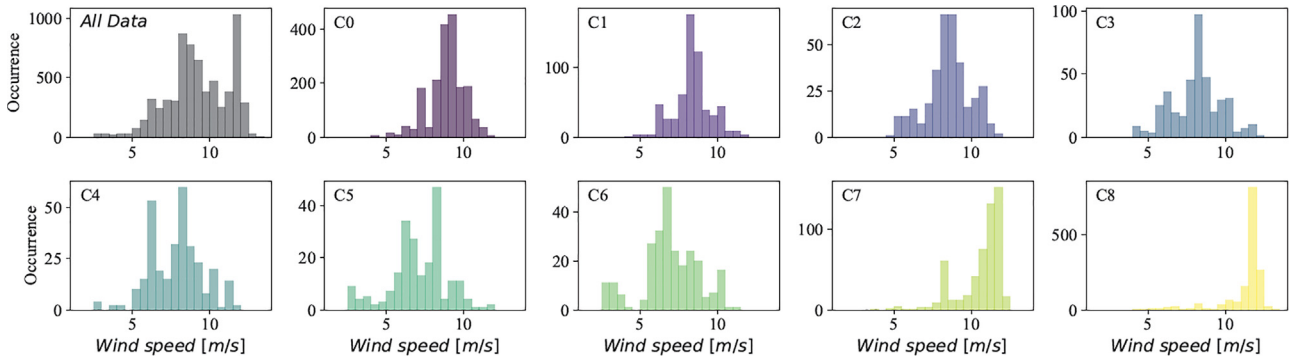


**FIG. 7.** Histograms of the hub-height wind speed recorded through the SCADA for the various clusters.

obtained from the statistics of the incoming-wind shear exponent, i.e., a larger value is reported for C0 (0.42), while the other clusters of region 2 can achieve a median value as low as 0.07.

The statistics of the meteorological parameters for the clusters owing to operations in the proximity of the rated wind speed, i.e., *C7* and *C8*, indicate the prevalence of neutral-stable conditions. Indeed, the median value of the shear exponent is larger than 0.3, the median *TI* is lower than 8% and the statistics of the Bulk Richardson number indicate predominantly stable atmospheric conditions for *C7* (median $Ri_B = 0.0018$) and neutral atmospheric conditions for *C8* (median $Ri_B = 0.0003$).

To conclude this statistical survey of the SCADA and meteorological data partitioned through the cluster analysis, we calculate the power coefficient as

$$C_P = \frac{Power}{\frac{1}{2}\rho U_{hub}^3 \pi \frac{d^2}{4}}, \tag{10}$$

where $U_{hub}$ is the hub-height mean wind speed recorded from the turbine anemometer through the SCADA, while the air density, $\rho$, is calculated from the air temperature and pressure measured from the met-tower. The minimum value of $C_P$ (0.26) is estimated for cluster *C8*, and it is slightly higher for *C7* (0.3) (see Table I). A reduction of $C_P$ with increasing wind speed in region 3 of the power curve is consistent with the operated blade pitching to keep power capture equal to the rater power of the turbine. The power coefficient, $C_P$ is larger for the clusters associated with operations in region 2 (*C0*–*C6*), with values noticeably larger for operations under unstable atmospheric conditions. This result corroborates a higher power efficiency of wind

**TABLE II.** Clustering of the SCADA and meteorological parameters. The superscript *p25*, *p50*, and *p75* indicate the 25th, 50th, and 75th percentiles. The horizontal blocks of the table (from top to bottom) are the statistics of the hub-height wind speed, power capture, pitch angle, Bulk Richardson number, hub-height wind turbulence intensity, and wind shear exponent.

| Parameter | All dataset | C0 | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 |
|---|---|---|---|---|---|---|---|---|---|---|
| $U_{hub}^{p25}$ (m/s) | 8.1 | 8.4 | 7.9 | 7.9 | 6.4 | 7.3 | 6.1 | 6.1 | 10.2 | 11.3 |
| $U_{hub}^{p50}$ (m/s) | 9.1 | 9.0 | 8.3 | 8.5 | 8.1 | 8.2 | 7.0 | 6.9 | 11.0 | 11.7 |
| $U_{hub}^{p75}$ (m/s) | 11.0 | 9.6 | 8.9 | 9.1 | 8.8 | 9.1 | 8.2 | 8.2 | 11.6 | 11.9 |
| $Power^{p25}$ (kW) | 1173 | 1233 | 1094 | 1098 | 664 | 874 | 591 | 600 | 1990 | 2265 |
| $Power^{p50}$ (kW) | 1530 | 1440 | 1279 | 1377 | 1173 | 1261 | 887 | 798 | 2239 | 2304 |
| $Power^{p75}$ (kW) | 2235 | 1680 | 1494 | 1690 | 1548 | 1652 | 1270 | 1216 | 2293 | 2309 |
| $Pitch^{p25}$ (deg) | −1.93 | −1.99 | −1.97 | −1.97 | −1.93 | −1.96 | −1.93 | −1.93 | −0.24 | 0.68 |
| $Pitch^{p50}$ (deg) | −1.70 | −1.93 | −1.91 | −1.88 | −1.83 | −1.80 | −1.83 | −1.82 | 1.26 | 2.64 |
| $Pitch^{p75}$ (deg) | 0.89 | −1.50 | −1.76 | −1.49 | −1.59 | −1.48 | −1.56 | −1.57 | 1.99 | 4.38 |
| $Ri_B^{p25}$ | −0.0019 | 0.0008 | −0.0021 | −0.0022 | −0.0032 | −0.0030 | −0.0044 | −0.0044 | −0.0001 | −0.0004 |
| $Ri_B^{p50}$ | −0.0003 | 0.0016 | −0.0014 | −0.0012 | −0.0022 | −0.0021 | −0.0034 | −0.0034 | 0.0018 | 0.0003 |
| $Ri_B^{p75}$ | 0.0018 | 0.0026 | −0.0007 | 0.0008 | −0.0015 | −0.0017 | −0.0026 | −0.0025 | 0.0038 | 0.0012 |
| $TI^{p25}$ (%) | 5.7 | 5.0 | 8.8 | 7.3 | 10.9 | 11.2 | 13.1 | 13.7 | 4.7 | 5.7 |
| $TI^{p50}$ (%) | 8.6 | 5.8 | 10.7 | 10.5 | 12.7 | 12.9 | 15.1 | 15.9 | 5.2 | 7.6 |
| $TI^{p75}$ (%) | 11.9 | 7.2 | 12.4 | 12.4 | 15.4 | 14.7 | 18.1 | 18.9 | 7.7 | 9.6 |
| $\alpha^{p25}$ | 0.09 | 0.16 | 0.05 | 0.07 | 0.04 | 0.04 | 0.04 | 0.05 | 0.21 | 0.20 |
| $\alpha^{p50}$ | 0.22 | 0.42 | 0.14 | 0.16 | 0.07 | 0.07 | 0.07 | 0.07 | 0.38 | 0.31 |
| $\alpha^{p75}$ | 0.39 | 0.52 | 0.33 | 0.35 | 0.13 | 0.16 | 0.12 | 0.12 | 0.45 | 0.37 |

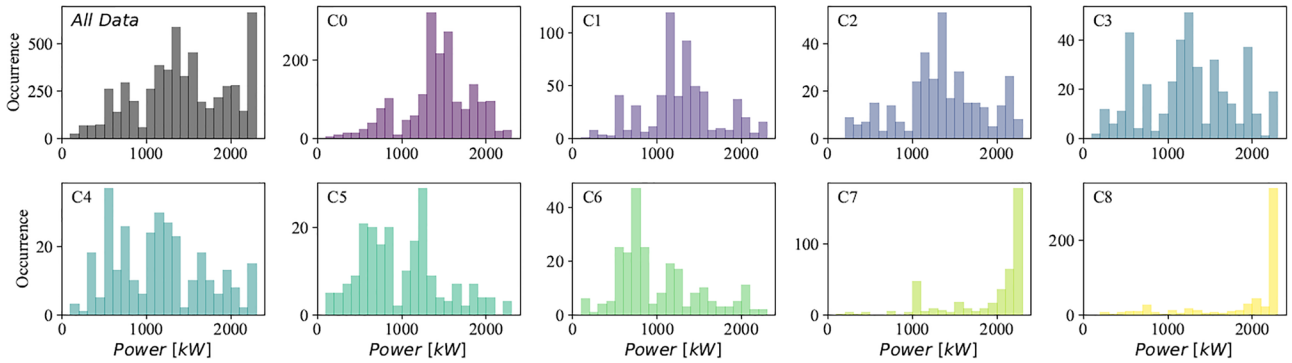


**FIG. 8.** Histograms of the power capture recorded through the SCADA for the various clusters.

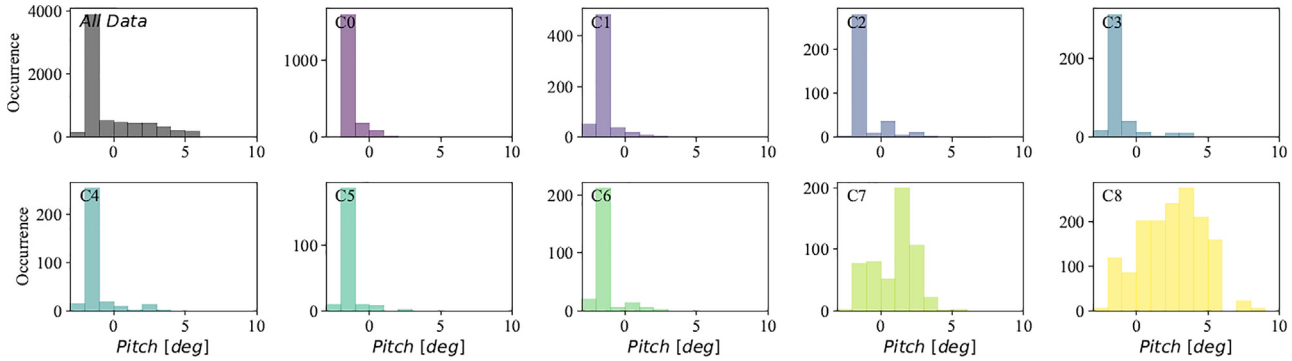


**FIG. 9.** Histograms of the blade pitch angle recorded through the SCADA for the various clusters.

turbines under convective conditions, as already observed through several wind farm studies.[3,74] However, it is noteworthy that the effect of atmospheric stability on power capture of wind turbines is still an open discussion, as highlighted from contrasting results obtained through numerical simulations,[22] or a lower power capture for convective atmospheric conditions estimated from other wind farm studies.[71,75]

By leveraging the actuator disk theory, it is possible to estimate the axial induction factor, $a$, from the power coefficient $[C_P = 4a(1-a)^2]$, and calculate, in turn, the thrust coefficient, $C_T^{SCADA} = 4a(1-a)$. The latter is reported in Table I for the various clusters, which confirms again an increasing axial induction of the turbine rotor moving from operations in region 3 to operations in region 2 under stable atmospheric conditions, and even further for operations in region 2 under

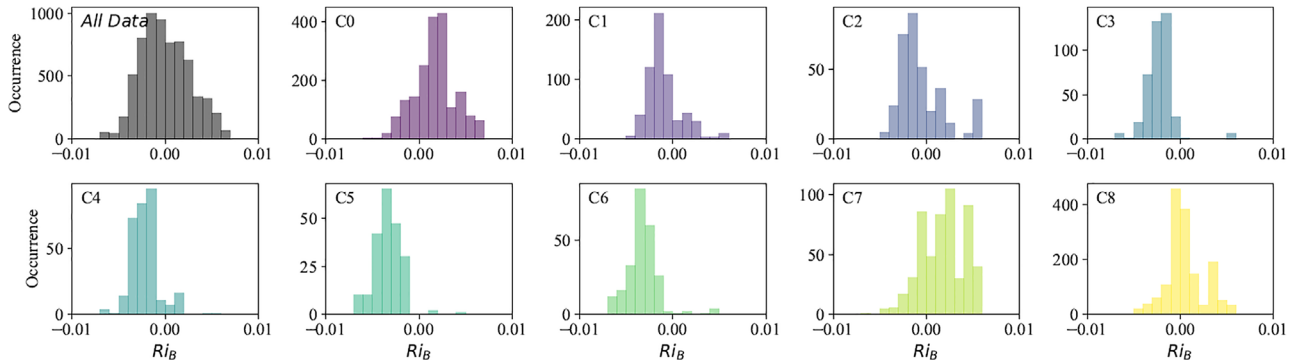


**FIG. 10.** Histograms of the Bulk Richardson number calculated from the meteorological data for the various clusters.

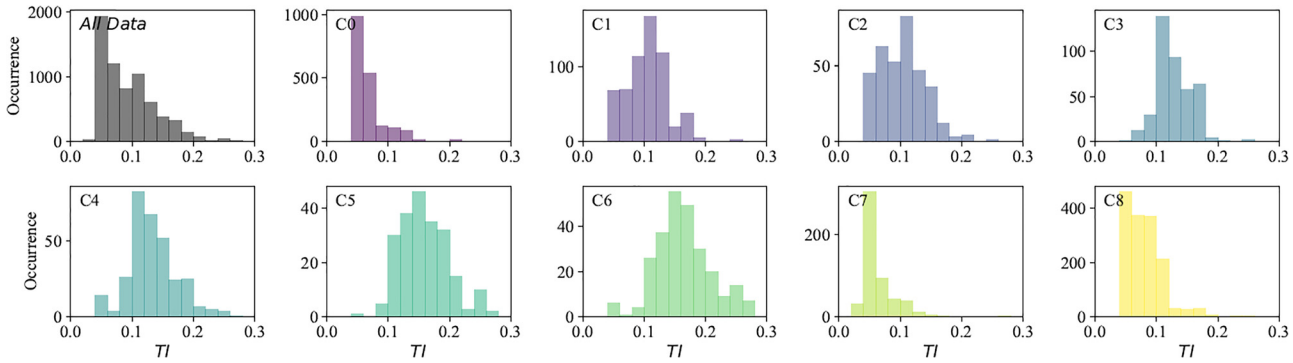


**FIG. 11.** Histograms of the hub-height turbulence intensity recorded through the SCADA for the various clusters.

unstable atmospheric conditions. More importantly, Table I shows that estimates of the rotor thrust coefficient derived from the power data, $C_T^{SCADA}$, are about half of the respective values obtained from the mass conservation and streamwise momentum budget applied to the mean LiDAR wake measurements with axisymmetry assumption, $C_T^{AD}$. This feature was already highlighted in previous works,[60] confirming that accurate estimates of the rotor thrust coefficient for wake modeling and predictions of power capture are still elusive.

Summarizing, this analysis of the SCADA and meteorological data for the various clusters, coupled with the analysis of the mean wake velocity field in Sec. V A, indicates that the clusters *C*0–*C*6 belong to operations in region 2 of the power curve, *C*7 to operations at the interface between region 2 and region 3 with active pitch control, and *C*8 to operations in region 3. Regarding the wake variability connected with atmospheric stability, *C*0 and *C*7 capture operations under stable conditions for regions 2 and 3 of the power curve, respectively. The clusters *C*1–*C*6 belong to operations under unstable conditions with a gradually increasing level of instability, while LiDAR data of *C*8 were collected under quasi-neutral conditions.

### C. Wake-added turbulence intensity for the various clusters

Wake turbulence intensity, $I_u = \sigma_u/U_\infty$, where $\sigma_u$ is the ensemble standard deviation of the streamwise velocity calculated through the cluster analysis (Sec. IV), affects directly the rotor aerodynamic forces ($\propto 1 + I_u^2$), and turbine power capture ($\propto 1 + 3I_u^2 + \gamma I_u^3$, where $\gamma$ is the skewness of the streamwise velocity).[76,77] Furthermore, turbulence intensity is very important in wind energy because it is typically used to model turbine fatigue loads.[78] Downstream of a turbine rotor, the wake-added standard deviation of the streamwise velocity, $\Delta u'$, is calculated from the standard deviation of the incoming wind velocity, $\sigma_{U_\infty}$, as[79]

$$\Delta u'(x,r) = sign\left[\sigma_u^2(x,r) - \sigma_{U_\infty}^2\right]\sqrt{|\sigma_u^2(x,r) - \sigma_{U_\infty}^2|}, \qquad (11)$$

where the operator *sign* produces values equal to 1 or −1 if $\sigma_u^2$ is larger or smaller than $\sigma_{U_\infty}^2$, respectively.

From hot-wire measurements carried out in a boundary-layer wind tunnel for the wake of a down-scaled wind turbine model, the wake-added turbulence intensity was observed to be either positive or negative over different wake areas and downstream locations.[77] In the far wake, statistics of the wake-added turbulence intensity, such as mean, maximum, minimum, and average of only positive or negative added turbulence intensity, showed the decay of these parameters in the downstream direction according to power laws with an exponent between −0.5 and −0.3. $\Delta u'$ reaches typically its maximum within the range $2 \lesssim x/d \lesssim 5$, which can coincide with the transition between the near- and far-wake regions.

A simple model for predicting the maximum of the wake-added turbulence intensity, $\Delta u'_{max}/U_\infty$, was proposed by Crespo and Hernández[80]

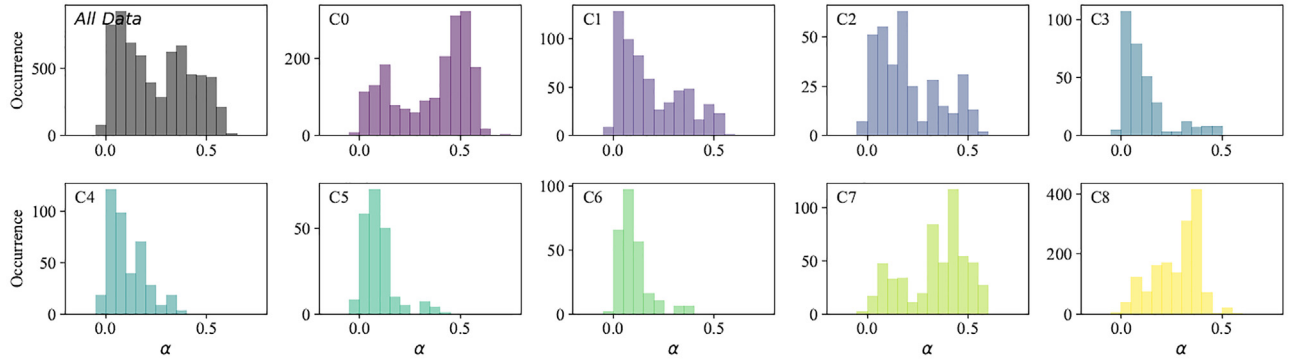


**FIG. 12.** Histograms of the shear exponent of the incoming wind for the various clusters.

$$\frac{\Delta u'_{max}}{U_\infty} = \begin{cases} 0.362\left(1-\sqrt{1-C_T}\right), & x/D \le 3, \\ 0.73\left(\dfrac{1-\sqrt{1-C_T}}{2}\right)^{0.83}, & x/D > 3, \end{cases} \tag{12}$$

while the following model[79] is used for the IEC61400–1 standard:[81]

$$\frac{\Delta u'_{max}}{U_\infty} = \frac{1}{1.5 + 0.8\, C_T^{-0.5} x/d}. \tag{13}$$

A similar model was proposed in Larsen, Hostrup, and Madsen[82]

$$\frac{\Delta u'_{max}}{U_\infty} = 0.29(x/d)^{-1/3}\sqrt{1-\sqrt{1-C_T}}. \tag{14}$$

In Quarton and Ainslie,[83] the downstream evolution of the peak turbulence intensity is modeled by using as the scaling parameter the location of the near-wake extent, $x_n$

$$\frac{\Delta u'_{max}}{U_\infty} = 4.8\, C_T^{0.7} I_u^{0.68} (x/x_n)^{-0.57}, \tag{15}$$

where $x_n$ is estimated as[7]

$$x_n = \frac{\sqrt{0.214+0.144m}\left(1-\sqrt{0.134+0.124m}\right)}{\left(1-\sqrt{0.214+0.144m}\right)\sqrt{0.134+0.124m}} \frac{r_0}{dr/dx}, \tag{16}$$

where $m = 1/\sqrt{10C_T}$ and $r_0 = 0.5d\sqrt{0.5(m+1)}$. The wake growth rate, $dr/dx$, can be estimated as

$$\frac{dr}{dx} = \sqrt{\left(\frac{dr}{dx}\right)_a^2 + \left(\frac{dr}{dx}\right)_m^2 + \left(\frac{dr}{dx}\right)_\lambda^2}, \tag{17}$$

where the three terms on the right-hand side represent the contribution due to ambient turbulence $[(dr/dx)_a^2 = 2.5\, I_u + 0.005]$, wake-generated turbulence $[(dr/dx)_m^2 = (1-m)\sqrt{1.49+m}/9.76/(1+m)]$, and mechanical turbulence $[(dr/dx)_\lambda^2 = 0.012 N_B \lambda$, where $N_B$ is the number of blades and $\lambda$ is the rotor tip-speed-ratio].

Another model to estimate the wake-added turbulence intensity was proposed by Hassan[84]

$$\frac{\Delta u'_{max}}{U_\infty} = 5.7\, C_T^{0.7} I_u^{0.68} (x/x_n)^{-0.96}. \tag{18}$$

For the analysis of the wake-added turbulence intensity for the various clusters, the freestream turbulence intensity, $\sigma_{U_\infty}/U_\infty$, is estimated from the ensemble standard deviation of the LiDAR measurements associated with each cluster, specifically from the median of $\sigma_u$ measured at $x/d = 1$ at sufficient lateral distance from the wake, typically $|r/d| > 0.7$. The respective values, which are reported in Table I, show variability among the various clusters similar to that observed for *TI* estimated from the SCADA data and reported in Table II. However, the values of $\sigma_{U_\infty}/U_\infty$ are significantly smaller than the respective ones estimated through the nacelle-mounted anemometers, which might be a consequence of the larger measurement volume of the LiDAR and the associated spatial averaging on the velocity fluctuations over each measurement volume.[85,86]

We analyze the statistics of the wake-added turbulence intensity in terms of minimum, maximum, and mean of $\Delta u'$; furthermore, mean values for the regions where $\Delta u'$ is only positive or negative are also calculated (referred to as $mean^+$ and $mean^-$, respectively). In Fig. 13, the statistics of $\Delta u'$ indicate that in the near wake, i.e., for $x/d \lesssim 3$, a region with turbulence intensity smaller than the incoming turbulence intensity ($\Delta u' < 0$) is generally observed for the clusters associated with operations in region 2 of the turbine power curve under unstable atmospheric conditions, i.e., for the clusters *C*1–*C*6. This wake feature can be associated with regularization of the flow induced by the wake swirl and the rotational flow induced by the near-wake vorticity structures, such as hub vortex and root vortices present at the core of the wake,[77,87–89] with respect to the surrounding ambient turbulence. By proceeding downstream, this region with negative $\Delta u'$ gradually fades out together with the diffusion of the mentioned wake vorticity structures.

In the near wake, the mean of $\Delta u'$ generally increases as a result of the shear-generated turbulence, which is due to the increasing velocity deficit associated with the thrust force induced by the turbine rotor. In the far-wake, $\Delta u'$ shows a significant reduction only for the clusters *C*5, *C*6, and *C*8. Indeed, the LiDAR velocity statistics in Fig. 5 show that the magnitude of the $\Delta u'$ peaks located roughly at the lateral boundaries of the wake reduce by proceeding downstream, which is a consequence of the wake recovery, while $\Delta u'$ increases at the wake core, which might be the footprint of far-wake dynamics, e.g., wake meandering.[50,90–92]

Predictions of the statistics for the wake-added turbulence intensity obtained through the various above-mentioned models [Eqs. (12)–(18)] are also reported in Fig. 13. These models generally produce realistic predictions of the maximum wake-generated turbulence; however, the downstream extent of the wake investigated ($1 \le x/d \le 7$) is far to show an exponential decay, as usually predicted through these models for the far-wake region. This is a critical limitation of the existing models for predicting the wake-added turbulence intensity because of the lack of accurate predictions for cases where wake interactions may occur with relatively small streamwise spacing among wind turbines, i.e., smaller than $7d$.[3]

## VI. CONCLUSIONS

LiDAR measurements of wind turbine wakes have been investigated through cluster analysis to identify the most representative wake morphologies associated with different atmospheric stability regimes, wind conditions, and control settings of the wind turbines. The wake LiDAR measurements are first projected on a truncated POD basis consisting of only five physics-informed POD modes. The reduced dimensionality of the experimental dataset has been instrumental to reduce the computational costs for the cluster analysis of two orders of magnitude.

The coefficients of the selected POD modes are then injected in a k-means algorithm, which identifies nine clusters to cover the variability of the wind turbine wakes observed through the LiDAR measurements. The synergistic analysis of the clustered LiDAR, meteorological, and SCADA data has enabled us to ascribe seven clusters to operations in region two of the power curve, namely, for incoming wind speeds lower than the turbine rated wind speed, and two clusters for operations above rated wind speed. While the latter mainly occur under neutral/stable atmospheric conditions, the other seven clusters owing to region two of the power curve are characterized by a varying level of atmospheric instability, leading to different incoming turbulence intensity and wake recovery rate.

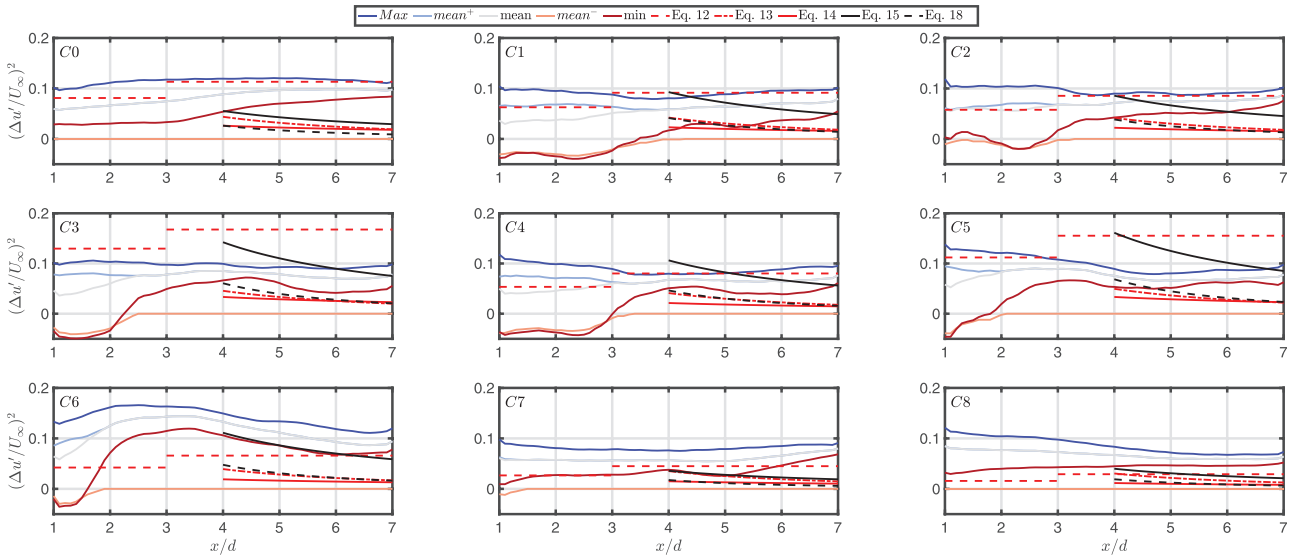


**FIG. 13.** Statistics of the wake-added turbulence intensity, $\Delta u'/U_\infty$, and comparison with the predictions obtained through several engineering models [Eqs. (12)–(18)].

It is noteworthy that the completely data-driven approach of the cluster analysis, which avoids the imposition of bounds for the various atmospheric, wind, and turbine parameters for detecting the wake variability, allows for the identification of systematic operations of the wind turbines with a certain yaw misalignment from the incoming mean wind direction. Indeed, for a duration of about 10% of the entire LiDAR experiment, the turbine rotors operated with a yaw angle between 3° and 5°, leading to significant deflections of the wind turbine wakes, and, eventually, effects on the wind farm power efficiency.

The clustered LiDAR data have also been analyzed in terms of wake-added turbulence intensity. Specifically, regions with a reduced wind turbulence intensity (negative wake-added turbulence intensity) have been observed in the near wake for operations in region two of the power curve and under unstable atmospheric conditions. This wake feature might be ascribed to the flow re-organization due to the wake vorticity structures, such as tip and hub vortices. In the far-wake, a general decay of the wake-added turbulence intensity is observed, yet not with a systematic exponential decay as proposed from existing wake models. However, predictions of the wake-added turbulence intensity obtained with several existing models have produced similar values to those obtained from the experimental data, albeit their accuracy is relatively poor in terms of streamwise variability. This analysis suggests that further work is needed for modeling the wake-added turbulence intensity for wake regions where the asymptotic exponential decay is not achieved yet, which can be an important flow feature in the presence of wake interactions with streamwise spacing smaller than about seven rotor diameters.

## ACKNOWLEDGMENTS

This research has been funded by from the National Science Foundation, Fluid Dynamics Program, Award No. 1705837, and the NSF CAREER program, Award No. 2046160. Pattern Energy Group is acknowledged to provide access to the wind farm for the LiDAR experiment and wind farm data. The Texas Advanced Computing Center is acknowledged for the computational resources. The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02–06CH11357. This work also utilized the resources of the Argonne Leadership Computing Facility under Contract No. DE-AC02–06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

## DATA AVAILABILITY

SCADA and meteorological data used in this work are protected by a non-disclosure agreement. LiDAR data can be requested to the corresponding author.

## REFERENCES

[1]R. J. Stevens and C. Meneveau, "Flow structure and turbulence in wind farms," Annu. Rev. Fluid Mech. **49**, 311–339 (2017).
[2]F. Porté-Agel, M. Bastankhah, and S. Shamsoddin, "Wind-turbine and wind-farm flows: A review," Boundary Layer Meteorol. **174**, 1–59 (2019).

[3]S. El-Asha, L. Zhan, and G. V. Iungo, "Quantification of power losses due to wind turbine wake interactions through SCADA, meteorological and wind LiDAR data," Wind Energy **20**, 1823–1839 (2017).
[4]M. B. Christiansen and C. B. Hasager, "Wake effects of large offshore wind farms identified from satellite SAR," Remote Sens. Environ. **98**, 251–268 (2005).
[5]R. Barthelmie, S. Pryor, S. Frandsen, K. Hansen, J. Schepers, K. Rados, W. Schelz, A. Neubert, L. Jensen, and S. Neckelmann, "Quantifying the impact of wind turbine wakes on power output at offshore wind farms," J. Atmos. Oceanic Technol. **27**, 1302–1317 (2010).
[6]A. Sebastiani, A. Segalini, F. Castellani, and G. Crasto, "Data analysis and simulation of the Lillgrund wind farm," Wind Energy **24**, 634–648 (2020).
[7]P. E. J. Vermeulen, "An experimental analysis of wind turbine wake," in Proceedings of the 3rd International Symposium on Wind Energy Systems, Copenhagen, August (1980).
[8]G. V. Iungo and F. Porté-Agel, "Volumetric LiDAR scanning of wind turbine wakes under convective and neutral atmospheric stability regimes," J. Atmos. Oceanic Technol. **31**, 2035–2048 (2014).
[9]L. Zhan, S. Letizia, and G. V. Iungo, "LiDAR measurements for an onshore wind farm: Wake variability for different incoming wind speeds and atmospheric stability regimes," Wind Energy **23**, 501–527 (2020).
[10]L. Zhan, S. Letizia, and G. V. Iungo, "Optimal tuning of engineering wake models through LiDAR measurements," Wind Energy Sci. **5**, 1601–1622 (2020).
[11]P. Veers, K. Dykes, E. Lantz, S. Barth, C. Bottasso, O. Carlson, A. Clifton, J. Green, P. Green, H. Holttinen, D. Laird, V. Lehtomäki, J. Lundquist, J. Manwell, M. Marquis, C. Meneveau, P. Moriarty, X. Munduate, M. Muskulus, J. Naughton, L. Pao, J. Paquette, J. Peinke, A. Robertson, J. S. Rodrigo, A. Sempreviva, J. Smith, A. Tuohy, and R. Wiser, "Grand challenges in the science of wind energy," Science **366**, eaau2027 (2019).
[12]C. Santoni, E. J. García-Cartagena, U. Ciri, L. Zhan, G. V. Iungo, and S. Leonardi, "One-way mesoscale-microscale coupling for simulating a wind farm in North Texas: Assessment against SCADA and LiDAR data," Wind Energy **23**, 691–710 (2020).
[13]C. Draxl, D. Allaerts, E. Quon, and M. Churchfield, "Coupling mesoscale budget components to large-eddy simulations for wind-energy applications," Boundary Layer Meteorol. **179**, 73–98 (2021).
[14]M. Calaf, C. Meneveau, and J. Meyers, "Large eddy simulation study of fully developed wind-turbine array boundary layers," Phys. Fluids **22**, 015110 (2010).
[15]F. Porté-Agel, Y.-T. Wu, H. Lu, and R. J. Conzemius, "Large-eddy simulation of atmospheric boundary layer flow through wind turbines and wind farms," J. Wind Eng. Ind. Aerodyn. **99**, 154–168 (2011).
[16]H. Lu and F. Porté-Agel, "Large-eddy simulation of a very large wind farm in a stable atmospheric boundary layer," Phys. Fluids **23**, 065101 (2011).
[17]C. Meneveau, "The top-down model of wind farm boundary layers and its applications," J. Turbul. **13**, N7 (2012).
[18]M. J. Churchfield, S. Lee, J, Michalakes, and P. J. Moriarty, "A numerical study of the effects of atmospheric and wake turbulence on wind turbine dynamics," J. Turbul. **13**, N14 (2012).
[19]C. L. Archer, S. Mirzaeisefat, and S. Lee, "Quantifying the sensitivity of wind farm performance to array layout options using large-eddy simulation," Geophys. Res. Let. **40**, 4963–4970, https://doi.org/10.1002/grl.50911 (2013).
[20]S. Xie and C. L. Archer, "Self-similarity and turbulence characteristics of wind turbine wakes via large-eddy simulation," Wind Energy **18**, 1815–1838 (2014).
[21]A. Creech, W.-G. Früh, and A. E. Maguire, "Simulations of an offshore wind farm using large-eddy simulation and a torque-controlled actuator disc model," Surv. Geophys. **36**, 427–481 (2015).
[22]N. S. Ghaisas, C. L. Archer, S. Xie, S. Wu, and E. Maguire, "Evaluation of layout and atmospheric stability effects in wind farms using large-eddy simulation," Wind Energy **20**, 1227–1240 (2017).
[23]S. Xie, C. L. Archer, N. Ghaisas, and C. Meneveau, "Benefits of collocating vertical-axis and horizontal-axis wind turbines in large wind farms," Wind Energy **20**, 45–62 (2016).
[24]M. Bastankhah and F. Porté-Agel, "A new miniature wind turbine for wind tunnel experiments. I. Design and performance," Energies **10**, 908 (2017).
[25]E. M. Nanos, C. L. Bottasso, F. Campagnolo, S. Letizia, G. V. Iungo, and M. A. Rotea, "Design, performance and wake characterization of a scaled wind turbine with closed-loop controls," Wind Energy Sci. Discuss. **2021**, 1–36.
[26]L. Neuhaus, M. Hölling, W. J. T. Bos, and J. Peinke, "Generation of atmospheric turbulence with unprecedentedly large Reynolds number in a wind tunnel," Phys. Rev. Lett. **125**, 154503 (2020).
[27]D. A. Rajewski1, E. S. Takle, J. K. Lundquist, S. Oncley, J. H. Prueger, T. W. Horst, M. E. Rhodes, R. Pfeiffer, J. L. Hatfield, K. K. Spoth, and R. K. Doorenbos, "Crop wind energy experiment (CWEX): Observations of surface-layer, boundary layer, and mesoscale interactions with a wind farm," Bull. Am. Meteorol. Soc. **94**, 655–672 (2013).
[28]C. L. Archer, S. Wu, A. Vasel-Be-Hagh, J. F. Brodie, R. Delgado, A. St. Pé, S. Oncley, and S. Semmer, "The VERTEX field campaign: Observations of near-ground effects of wind turbine wakes," J. Turbul. **20**, 64–92 (2019).
[29]N. Bodini, J. K. Lundquist, and A. Kirincich, "Offshore wind turbines will encounter very low atmospheric turbulence," J. Phys.: Conf. Ser. **1452**, 012023 (2020).
[30]S. Wu and C. L. Archer, "Near-ground effects of wind turbines: Observations and physical mechanisms," Mon. Weather Rev. **149**, 879–898 (2021).
[31]G. V. Iungo, Y.-T. Wu, and F. Porté-Agel, "Field measurements of wind turbine wakes with lidars," J. Atmos. Oceanic Technol. **30**, 274–287 (2013).
[32]B. D. Hirth, J. L. Schroeder, W. S. Gunter, and J. G. Guynes, "Coupling Doppler radar derived wind maps with operational turbine data to document wind farm complex flows," Wind Energy **18**, 529–540 (2015).
[33]G. Kocer, M. Mansour, N. Chokani, R. Abhari, and M. Müller, "Full-scale wind turbine near-wake measurements using an instrumented uninhabited aerial vehicle," J. Sol. Energy Eng. **133**, 041011 (2011).
[34]A. Platis, S. K. Siedersleben, J. Bange, A. Lampert, K. Bärfuss, R. Hankers, B. Cañadillas, R. Foreman, J. Schulz-Stellenfleth, B. Djath, T. Neumann, and S. Emeis, "First in situ evidence of wakes in the far field behind offshore wind farms," Sci. Rep. **8**, 2163 (2018).
[35]S. Letizia, L. Zhan, and G. Valerio Iungo, "LiSBOA: LiDAR statistical Barnes objective analysis for optimal design of LiDAR scans and retrieval of wind statistics. I. Theoretical framework," Atmos. Meas. Tech. **14**, 2065–2093 (2021).
[36]S. Letizia, L. Zhan, and G. Valerio Iungo, "LiSBOA: LiDAR statistical Barnes objective analysis for optimal design of LiDAR scans and retrieval of wind statistics. II. Applications LiDAR measurements of wind turbine wakes," Atmos. Meas. Tech. **14**, 2095–2113 (2021).
[37]See https://www.usgs.gov/ for U.S. Geological Survey Website; accessed 06 August 2021.
[38]R. B. Stull, *An Introduction to Boundary Layer Meteorology* (Springer, 1988).
[39]J. D'Errico, see https://www.mathworks.com/matlabcentral/fileexchange/4551-inpaint_nans for "Inpaint-nans, MATLAB Central File Exchange," (2004) (last accessed 06 August 2021).
[40]H. Li, D. Fernex, R. Semaan, J. Tan, M. Morzyński, and B. R. Noack, "Cluster-based network model," J. Fluid Mech. **906**, A21 (2020).
[41]G. V. Iungo and E. Lombardi, "Time-frequency analysis of the dynamics of different vorticity structures generated from a finite-length triangular prism," J. Wind Eng. Ind. Aerodyn. **99**, 711–717 (2011).
[42]G. V. Iungo and E. Lombardi, "A procedure based on proper orthogonal decomposition for time-frequency analysis of time series," Exp. Fluids **51**, 969–985 (2011).
[43]M. Debnath, C. Santoni, S. Leonardi, and G. V. Iungo, "Towards reduced order modelling for predicting the dynamics of coherent vorticity structures within wind turbine wakes," Philos. Trans. R. Soc. A **375**, 20160108 (2017).
[44]J. L. Lumley, *Stochastic Tools in Turbulence* (Academic Press, New York, NY, 1970).
[45]G. Berkooz, P. Holmes, and J. L. Lumley, "The proper orthogonal decomposition in the analysis of turbulent flows," Annu. Rev. Fluid Mech. **25**, 539–575 (1993).
[46]P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry* (Cambridge University Press, Cambridge, UK, 1996).
[47]L. Sirovich, "Turbulence and the dynamics of coherent structures," Q. Appl. Math. **45**, 561–571 (1987).
[48]G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning* (Springer, 2013).
[49]A. Abraham and J. Hong, "Characterization of atmospheric coherent structures and their impact on a utility-scale wind turbine," arXiv:2110.06279 (2021).

[50]D. Foti, X. Yang, M. Guala, and F. Sotiropoulos, "Wake meandering statistics of a model wind turbine: Insights gained by large eddy simulations," Phys. Rev. Fluids **1**, 044407 (2016).
[51]M. Abkar and F. Porté-Agel, "Influence of the Coriolis force on the structure and evolution of wind turbine wakes," Phys. Rev. Fluids **1**, 063701 (2016).
[52]A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," ACM Comput. Surv. **31**, 264–323 (1999).
[53]A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," Pattern Recognit. **36**, 451–461 (2003).
[54]P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," J. Compt. Appl. Math. **20**, 53–65 (1987).
[55]M. Forina, C. Armanino, and V. Raggio, "Clustering with dendrograms on interpretation variables," Anal. Chim. Acta **454**, 13–19 (2002).
[56]M. Bastankhah and F. Porté-Agel, "Experimental and theoretical study of wind turbine wakes in yawed conditions," J. Fluid Mech. **806**, 506–541 (2016).
[57]J. F. Ainslie, "Calculating the flow field in the wake of wind turbines," J. Wind. Eng. Ind. Aerodyn. **27**, 213–224 (1988).
[58]J. N. Sørensen, R. F. Mikkelsen, D. S. Henningson, S. Ivanell, S. Sarmast, and S. J. Andersen, "Simulation of wind turbine wakes using the actuator line technique," Philos. Trans. R. Soc. London, Ser. A **373**, 20140071 (2015).
[59]H. A. Madsen, G. C. Larsen, T. J. Larsen, N. Troldborg, and R. Mikkelsen, "Calibration and validation of the dynamic wake meandering model for implementation in an aeroelastic code," J. Sol. Energy Eng. **132**, 041014 (2010).
[60]G. V. Iungo, S. Letizia, and L. Zhan, "Quantification of the axial induction exerted by utility-scale wind turbines by coupling LiDAR measurements and RANS simulations," J. Phys.: Conf. Ser. **1037**, 072023 (2018).
[61]G. K. Batchelor, *An Introduction to Fluid Mechanics* (Cambridge University Press, Cambridge, UK, 1967).
[62]S. B. Pope, *Turbulent Flows* (Cambridge University Press, Cambridge, UK, 2000).
[63]See https://github.com/UTD-WindFluX/G-RANS for "WindFluX Global 3D RANS code for simulations of wind turbine wakes," (last accessed 06 August 2021).
[64]R. Fletcher, *Practical Methods of Optimization* (John Wiley and Sons, Inc., 1987).
[65]M. Howland, L. A. Bossuyt, J. Martinez-Tossas, J. Meyers, and C. Meneveau, "Wake structure in actuator disk models of wind turbines in yaw under uniform inflow conditions," J. Renewable Sustainable Energy **8**, 043301 (2016).
[66]P. M. O. Gebraad, F. W. Teeuwisse, J. W. van Wingerden, P. A. Flemming, S. D. Ruben, J. R. Marden, and L. Y. Pao, "Wind plant power optimization through yaw control using a parametric model for wake effects—A CFD simulation study," Wind Energy **19**, 95–114 (2016).
[67]M. Bastankhah and F. Porté-Agel, "A new analytical model for wind-turbine wakes," Renewable Energy **70**, 116–123 (2014).
[68]T. Burton, N. Jenkins, D. Sharpe, and E. Bossanyi, *Wind Energy Handbook* (John Wiley and Sons, Inc., 2011).
[69]M. Magnusson and A. S. Smedman, "Influence of atmospheric stability on wind turbine wakes," Wind Eng. **18**, 139–152 (1994).
[70]K. S. Hansen, R. J. Barthelmie, L. E. Jensen, and A. Sommer, "The impact of turbulence intensity and atmospheric stability on power deficits due to wind turbine wakes at Horns Rev wind farm," Wind Energy **15**, 183–196 (2015).
[71]S. Wharton and J. K. Lundquist, "Atmospheric stability affects wind turbine power collection," Environ. Res. Lett. **7**, 014005 (2012).
[72]M. Abkar and F. Porté-Agel, "Influence of atmospheric stability on wind-turbine wakes: A large-eddy simulation study," Phys. Fluids **27**, 035104 (2015).
[73]S. Xie and C. L. Archer, "A numerical study of wind-turbine wakes for three atmospheric stability conditions," Boundary Layer Meteorol. **165**, 87–112 (2017).
[74]B. J. Vaderwende and J. K. Lundquist, "The modification of wind turbine performance by statistically distinct atmospheric regimes," Environ. Res. Lett. **7**, 034035 (2012).
[75]S. Wharton and J. K. Lundquist, "Assessing atmospheric stability and its impacts on rotor-disk wind characteristics at an onshore wind farm," Wind Energy **15**, 525–546 (2012).
[76]L. Binh, T. Ishihara, P. van Phuc, and Y. Fujino, "A peak factor for non-Gaussian response analysis of wind turbine tower," J. Wind Eng. Ind. Aerodyn. **96**, 2217–2227 (2008).
[77]L. Chamorro and F. Porté-Agel, "A wind-tunnel investigation of wind-turbine wakes: Boundary-layer turbulence effects," Boundary Layer Meteorol. **132**, 129–149 (2009).
[78]A. Rosen and Y. Sheinman, "The power fluctuations of a wind turbine," J. Wind Eng. Ind. Aerodyn. **59**, 51–68 (1996).
[79]S. Frandsen, "Turbulence and turbulence-generated structural loading in wind turbine clusters," Ph.D. thesis (Risoe National Laboratory, Wind Energy Department, Danmarks Tekniske University, Kgs. Lyngby, Roskilde, Denmark, 2007), No. rISO-R-1188.
[80]A. Crespo and J. Hernández, "Turbulence characteristics in wind-turbine wakes," J. Wind Eng. Ind. Aerodyn. **61**, 71–85 (1996).
[81]IEC 61400-12-1-2015, "Ed 2.0: Wind turbines. 12-1. Power performance measurements of electricity producing wind turbines," Technical Report No. IEC 61400-12-1-2015 (IEC, Geneva, Switzerland, 2015).
[82]G. C. Larsen, J. Hostrup, and H. Madsen, "Wind fields in wakes," in *Proceedings of the European Union Wind Energy Conference*, Bedford (Bedford: Stephens & Associates, 1996), pp. 764–768.
[83]D. C. Quarton and J. F. Ainslie, "Turbulence in wind turbine wakes," Wind Eng. **14**, 15–23 (1990); available at https://www.jstor.org/stable/43749409?seq=1.
[84]U. Hassan, "A wind tunnel investigation of the wake structure within small wind turbine farms," Technical Report No. ETSU-WN-5113 (UK Department of Energy, 1992).
[85]P. Brugger, K. Träumner, and C. Jung, "Evaluation of a procedure to correct spatial averaging in turbulence statistics from a Doppler LiDAR by comparing time series with an ultrasonic anemometer," J. Atmos. Oceanic Technol. **96**, 2135–2144 (2016).
[86]M. Puccioni and G. V. Iungo, "Spectral correction of turbulent energy damping on wind LiDAR measurements due to spatial averaging," Atmos. Meas. Tech. **14**, 1457–1474 (2021).
[87]G. V. Iungo, F. Viola, S. Camarri, F. Porté-Agel, and F. Gallaire, "Linear stability analysis of wind turbine wakes performed on wind tunnel measurements," J. Fluid Mech. **737**, 499–526 (2013).
[88]F. Viola, G. V. Iungo, S. Camarri, F. Porté-Agel, and F. Gallaire, "Prediction of the hub vortex instability in a wind turbine wake: Stability analysis with eddy-viscosity models calibrated on wind tunnel data," J. Fluid Mech. **750**, R1 (2014).
[89]R. Ashton, F. Viola, S. Camarri, F. Gallaire, and G. V. Iungo, "Hub vortex instability within wind turbine wakes: Effects of wind turbulence, loading conditions, and blade aerodynamics," Phys. Rev. Fluids **1**, 073603 (2016).
[90]E. Keck, M. de Maré, M. Churchfield, S. Lee, G. Larsen, and H. A. Madsen, "On atmospheric stability in the dynamic wake meandering model," Wind Energy **17**, 1689–1710 (2014).
[91]G. C. Larsen, E. Machefaux, and A. Chougule, "Wake meandering under non-neutral atmospheric stability conditions—Theory and facts," J. Phys.: Conf. Ser. **625**, 012036IOP (2015).
[92]Y. A. Muller, S. Aubrun, and C. Masson, "Determination of real-time predictors of the wind turbine wake meandering," Exp. Fluids **56**, 53 (2015).
[93]R. Maulik, V. Rao, S. A. Renganathan, S. Letizia, and G. V. Iungo, "Cluster analysis of wind turbine wakes measured through a scanning Doppler wind LiDAR," in *AIAA Scitech 2021 Forum* (AIAA, 2021), pp. 2021–1181.