Wind turbine wake detection and characterisation utilising blade loads and SCADA data: a generalised approach

Piotr Fojcik[1], Edward Hart[1], and Emil Hedevang[2]

[1]Electronic & Electrical Engineering, University of Strathclyde, 99 George Street, Glasgow G1 1RD, United Kingdom
[2]Siemens Gamesa Renewable Energy, Borupvej 16, 7330 Brande, Denmark

Correspondence: Piotr Fojcik (piotr.fojcik@strath.ac.uk)

Abstract. Large offshore wind farms face operational challenges due to turbine wakes, which can reduce energy yield and increase structural fatigue. These problems may be mitigated through wind farm flow control techniques, which require reliable wake detection (recognising the presence of a clear wake) and characterisation (parametric description of a wake's properties) as prerequisites. This paper presents a novel three-stage framework for generalised wake detection and characterisation. First, a regression model utilises blade loads and SCADA data to estimate the wind speed distribution across the rotor plane. Second, a ~~Convolutional Neural Network (CNN)~~ convolutional neural network undertakes pattern recognition analysis to perform the wake detection, classifying rotor-plane wind estimates as ~~"~~'fully-impinged~~", "~~', 'left-impinged~~", "~~', 'right-impinged~~" or "~~' or 'not impinged.~~"~~' Third, where wake impingement is detected, ~~2D~~ two-dimensional Gaussian fitting is undertaken to provide a parametric wake characterisation, providing outputs of the wake centre location and wake lateral width. The framework is tested ~~and assessed using a virtual wind farm in the North Sea and~~ in a simulation environment incorporating the Mann turbulence model, DWM model for generating wakes and BHawC aeroelastic code. The testing is undertaken under a wide range of wind conditions~~(~~, with mean ambient wind speeds from 5-15 m/s, turbulence intensities from 3-9% ~~,~~ and full range of wind directions~~)~~. Results show high accuracy of wind field estimation, with the mean RMSE over all test cases being ~~0.351 m/s, or 5.23~~3.7% when normalised by mean ambient wind speed. A wake detection sensitivity study confirms accurate performance across a majority of wind conditions, with minor issues observed only for more extreme conditions or those at the limits of the utilised training data. The final wake characterisation stage is shown to flexibly adapt to changing wind conditions, successfully tracking the wake's position even ~~in more demanding partial-impingement cases~~for more turbulent conditions. The proposed framework therefore demonstrates strong potential as a generalised approach to wake detection and characterisation.

1 Introduction

The wind industry has advanced significantly in recent years, reaching a total worldwide installed capacity of 906 GW at the ~~dawn of 2023~~ end of 2022 (Hutchinson and Zhao, 2023). Despite being a mature field, some challenges still remain to be solved. Due to the increasing size of modern wind farms, there is a newly-posed challenge of effectively operating clusters of hundreds of multi-megawatt machines. For this reason, the current focus for both research and industry is shifting towards a farm-level approach for operations, control and maintenance.

A wind turbine interacts with the incoming wind flow, extracting energy and creating a wake: a downstream region of decreased wind speed and increased turbulence. Due to their varied impact on the farm's performance, the ~~wind turbine wakes are an important aspect of the aforementioned task. First of all, operating a wind turbine affected by wake deficits from the upstream machines results in~~ wakes play a crucial role in turbine operation. A turbine affected by wakes from upstream machines has a lower energy yield ~~because of~~ due to a wake-generated velocity deficit in the upcoming flow (Adaramola and Krogstad, 2011; Barthelmie et al., 2010). Having the rotor experience ~~an altered~~ a more turbulent wind field also leads to uneven aerodynamic loading on the machine, inducing severe fatigue cycles ~~to~~ in its structure (Churchfield et al., 2012). A turbine under waked flow conditions is also prone to experiencing additional turbulence due to large-scale motions of wake meandering, creating substantial loads and thus fatigue (Madsen et al., 2010; Larsen et al., 2013).

Via various control techniques it is possible to mitigate the negatives of waked flow conditions, optimizing turbine operation for reliable and effective performance. Probably the earliest developments in this area date back to a study by Steinbuch et al. (1988), where the authors ~~implemented~~ implement axial induction control to increase the energy yield by reducing the wake effects. ~~Since then, numerous~~ Numerous methods for optimizing the wind farm flow ~~have been developed~~follow; some examples include ~~yaw control towards directing the wake away from the downstream machine~~ wake steering by introducing yaw offsets (Howland et al., 2020; SGRE, 2019), or ~~cyclic pitch control towards inducing improved~~ dynamic individual pitch control (Frederik et al., 2020) and dynamic induction control (Munters and Meyers, 2018) to induce enhanced mixing in the wake~~(Frederik et al., 2020)~~.

To use the above-mentioned approaches in a closed-loop control scheme, dynamic information on whether the impinging wake is being successfully redirected or dispersed is required (Raach et al., 2016). Moreover, before starting the control action, there first needs to be a confirmation that a turbine is indeed wake-affected to facilitate an intervention to its normal operating cycle. For the sake of the discussion in this work, we ~~introduce two terms:~~ term these two flow control prerequisites *wake detection* and *wake characterisation*. By the former we refer to the action of recognising that a given turbine is experiencing a clear wake impingement from a nearby machine, in such a way that it has a significant (in both magnitude and time) influence on its performance. By the latter we refer to the action of identifying the properties of the wake using a simplified parametric representation, and monitoring how the values of these parameters change in time. ~~These two steps are key prerequisites to wind farm flow control, which is why lately there has been a noticeable~~

There is a growing effort in the research community ~~towards developing methods to obtain detailed information on the instantaneous wind field impinging downstream turbines. So far, these approaches have focused~~ to develop new wake detection and characterisation techniques. Up-to-date approaches focus on implementing remote sensing techniques or using a turbine's operational data. The former is usually achieved with the use of ~~LIDAR devices - some examples include works by Lio et al. (2020) and Raach et al. (2016). A different way of tackling the problem is to employ the relationship between wind field and turbine response (primarily blade-root bending moments), which allows for a comparable accuracy in wake detection and characterisation (Onnen et al., 2022; Dong et al., 2021; Farrell et al., 2022).~~

~~Despite achieving good wake tracking~~ Light Detection and Ranging (LIDAR) devices, with recent literature featuring numerous examples of their successful application in wake characterisation (Bingöl et al., 2010; Trujillo et al., 2011; Conti et al., 2020; Lio

. A recent study (Raach et al., 2016) investigates the use of LIDAR-based wake characterisation directly in a closed-loop wake steering scheme, showing promising results. Despite achieving great performance, the LIDAR-based approaches unavoidably rely on an additional hardware, which is not currently present at most wind farms and represents a significant additional cost. ~~As for the~~ A different way of tackling the problem is to employ the relationship between wind field and turbine response. The basic idea behind this approach is that a variation in the incoming flow can be correlated with changes in signals such as blade root bending moments and generator speed. The widely discussed method introduced by Bottasso and Schreiber (2018) relies on using the out-of-plane blade root bending moments to approximate the local wind speed in each of the rotor quadrants, thus allowing for partial wake detection. The method is tested in the field (Schreiber et al., 2020) using a setup of two wind turbines and a met mast, providing a proof of concept for the qualitative wake detection. The idea of mapping between the blade loads and local wind speeds is also used by others, some examples include studies by Kim et al. (2023), Simley and Pao (2016) and Liu et al. (2021). A study by Onnen et al. (2022) pushes this approach further, implementing both in-plane and out-of-plane blade loads to achieve accurate characterisation of the meandering wake deficit with the use of Extended Kalman Filter (EKF). The authors also discuss the potential of using the estimation uncertainty metric to detect the wake presence. The method is validated via wind tunnel (Onnen et al., 2023) and field (Onnen et al., 2025) experiments, showing a strong correlation between the estimated wake position and the reference LIDAR measurements. A similar blade-load-based approach is also considered by Dong et al. (2021), where the authors investigate wake characterisation with an EKF setup within a simulation environment. In another study by Farrell et al. (2022), the authors implement a recurrent neural network trained with experimental and simulation data to estimate the lateral position of the wake centre.

All in all, we identify the research gap as follows: the existing wake estimation methods primarily focus on the wake characterisation aspect, with both LIDAR- and load-based ~~methods, to-date the studies focused solely on a scenario of~~ approaches offering high accuracy. The up-to-date wake detection studies analysed the wake impingement in a scenario with a single upwind turbine, for a fixed set of wind directions. This handles only a ~~direct and guaranteed wake impingement from an upstream turbine. This therefore only handles a~~ part of the overall problem in a realistic setting, ~~since for the majority~~ where the turbine experiences a full range of wind directions~~the turbine will not be~~. For the majority of inflow angles the turbine is not directly affected by a nearby ~~turbines~~ turbine’s wake, it ~~will~~ does however receive a large amount of random flow fluctuations due to operating in the highly turbulent atmospheric boundary layer. Without ~~knowing~~ a robust knowledge of whether the turbine operates under wake impingement or not, there is a risk of implementing a wake characterisation scheme on a naturally occurring turbulent eddy, hence providing incorrect and misleading inputs to the control system. ~~For this reason, the vast majority of methods developed so far are not yet applicable to real world operations. Therefore, to the best of the authors’ knowledge, there has not previously been a generalised method propoposed, which allows for both wake detection and a subsequent wake characterisation (if the flow field is recognised as being wake-impinged).~~

In this paper, we seek to bridge the above identified gap by proposing a generalised 3-stage approach that performs ~~both~~ wind field reconstruction, wake detection and wake characterisation within a single framework. Firstly, the instantaneous wind field interacting with the rotor is estimated using turbine operational and load data. Secondly, a ~~Neural Network~~ neural network approach is implemented for classifying whether the estimated wind field represents a case of ~~'~~‘full’, ~~'~~‘partial’ or ~~'~~‘no’ wake impingement. Thirdly, if clear wake impingement is detected, a parametric wake model is fitted in order to characterise the centre and lateral-width of the currently observed wake. A ~~full~~ novel end-to-end methodology is presented, which aims to provide both a demonstrator and a performance benchmark for generalised wake detection and characterisation methods of this type. ~~It is highlighted that focus of the current work is that of developing a solution which is able to~~ The models are trained and tested for a full range of wind directions within a virtual offshore wind farm. The simulation environment incorporates Mann turbulence wind boxes, Dynamic Wake Meandering (DWM) model for generating wake interactions, and aeroelastic code to compute the turbine response. This work focuses on developing a novel solution that can confidently assert when a turbine is impinged by a wake from a nearby ~~turbine, as this information is critical to farm level~~ machine - this being a key factor for farm-wide wake steering control.

The paper is structured as follows: Sect. ~~?? introduces core concepts which are key for understanding the scope of this work and provides a literature review of state-of-the-art in wind field reconstruction and wake estimation techniques; Sect.~~ 2 discusses the methodology used for the framework's development and training; Sect. 3 presents the results of the framework validation; Sect. 4 considers the discussion on the framework applicability and limitations; and lastly, Sect. 5 covers the conclusion and addresses future work.

2 ~~Background~~

~~This section describes the concepts of Atmospheric Boundary Layer, physics of turbine wake and wind farm flow as well as numerical turbine/farm flow models. It also covers the state-of-the-art in wake estimation and explains some of the methods used in the scope of this work.~~

1.1 ~~Atmospheric Boundary Layer~~

~~The turbines operate in the atmospheric boundary layer (ABL), a highly turbulent environment where all motions are affected by the proximity of the Earth's surface. Its height, usually associated with a characteristic value of 1 km (Calaf et al., 2010), can vary depending on the weather conditions and diurnal cycle. It introduces complex gradients of heat, momentum and humidity that strongly influence the wind farm's performance (Stevens and Meneveau, 2017). The general behaviour of the ABL is often described with its stability, which is related to the vertical temperature distribution and can be one of three types: *unstable* conditions (increased vertical fluctuations, horizontal flow turbulent), *stable* conditions (reduced vertical fluctuations, horizontal flow near-laminar) and *neutral* conditions (insignificant effect of vertical movements to the horizontal flow) (Keck et al., 2014). The terms *wind shear* and *wind veer* refer to changes in wind speed and direction with increasing height, respectively. The turbulent flow structures - termed *eddies* - are present in the ABL at different time and space scales. Most relevant for this project is the microscale, covering changes occurring over seconds or minutes with the length scales of up to 1 km. The most common way to describe the severity of the atmospheric fluctuations is the parameter of ambient turbulence intensity $I_{amb} = \sigma_U / \overline{U}$, where σ_U is the standard deviation of the stream-wise velocity fluctuations and $\overline{U}$ is the average stream-wise velocity.~~

1.1 ~~Wind turbine wake physics~~

~~A significant effort was put in by the scientific community to determine how a wind turbine interacts with the incoming flow. Thanks to decades of numerical, wind tunnel and field studies (Ronsten, 1992; Whale et al., 2000; Barthelmie et al., 2004; Medici and Alfre this topic is well understood today. The recent literature is also enriched with detailed review articles (Stevens and Meneveau, 2017; Porté-A covering numerous aspects of the wind turbine and wind farm flow.~~

~~A diagram showing how the rotor area A, mean wind speed U and mean pressure p change due to the flow interacting with the rotor. Rotor plane presented with a dotted grey line.~~

~~The actuator disc concept (Froude, 1889), being the cornerstone of most analytical techniques used today, is a reasonable simplification of the flow transitions happening at the wind turbine and is thus a proper introduction into any discussion on the wake subject. It considers the rotor as an infinitely thin permeable disc, inducing an axisymmetric force on the passing flow. As the flow approaches the moving blades, it is being slowed by the resistance of the rotor. As shown in the Figure **??**, the initial velocity value of U_0 is gradually reduced to a value of $U_0(1-a)$ at the rotor plane. The value of a is referred to as the axial induction factor. To satisfy the mass flow conservation condition, decreasing velocity results in the cross-sectional area A of the flow increasing. The energy extraction occurs only at the rotor plane, so before the flow reaches that point, its energy must remain unchanged. At the rotor plane, the extraction of energy manifests as a sudden drop in pressure. In the downwind region, the velocity keeps decreasing alongside a gradual increase in pressure until reaching the initial (atmospheric) level of p_0. The energy extraction is now visible in the velocity deficit (U_2-U_0). From the above discussion it should be clear that: a) kinetic energy is extracted from the flow at the global level (not locally); b) this is driven by the turbine extracting pressure (potential) energy at the local level.~~

~~The downstream region of the stream-wise flow is referred to as *wake*. It can be divided into the near-wake, considered to be up to 2-4 rotor diameters from the rotor, and the far-wake, spanning further downstream. These two flow structures are significantly different. The well-structured near-wake is directly linked to the rotor properties, manifested with tip and hub vortices. This region is of secondary importance for this discussion, as the modern turbines are spaced at such distances that they usually are in the far-wake of their upstream neighbours (Meyers et al., 2022). In the far-wake region, with the organised vortices dissipated, the flow is able to recover its kinetic energy by entrainment of the surrounding flow. The velocity deficit is gradually minimised, while the wake width increases. For the sake of simplicity, in this work the term 'wake' refers to the far-wake flow region.~~

~~In an idealised approach, where the wind field is simplified to a lateral flow with no ground present, mean velocity profile in a wake follows a near-perfect~~[1] ~~Gaussian distribution; also, with x considered to be downstream distance, the velocity deficit value at the rotor centre decreases proportionally to $x^{-2/3}$, while the wake width grows proportionally to $x^{1/3}$. The non-disturbed mean velocity profiles are characterised with self-similarity, meaning that the deficit shape (normalized by its maximum value), considered as a function of the radial distance from wake centre (normalized by wake width), remains the same along the streamwise direction.~~

---

[1] ~~There are corrections for e.g. hub and blade tips.~~

~~The evolution of mean wind speed profiles with the ABL flow moving in the streamwise direction x; shown as (a) function of vertical distance z (view from the side), (b) function of horizontal distance y (view from above). Near-wake profiles omitted due to not being the focus of this work. Adapted from (Stevens and Meneveau, 2017).~~

~~With this baseline explained, one should consider the wake physics in a more realistic setting. Figure **??** shows the mean U profiles typical for the ABL wind field and how they are altered after interacting with a rotor. Firstly focusing on the flow velocity, it can be seen that the stream-wise wake development defers from the idealised scenario mentioned previously. First of all, due to the atmospheric wind shear effect, the mean velocity profile is skewed towards the ground, resulting in a higher value in the upper part of the rotor. In (Stevens and Meneveau, 2017) the authors present this effect by comparing several different numerical and experimental studies. Furthermore, extensive numerical (Peña and Rathmann, 2014), wind tunnel (Bastankhah and Porté-Agel, 2014) and field (Fuertes et al., 2018) studies have shown that the atmospheric turbulence causes the wake width to grow approximately linearly with increasing downstream distance. A more rapid wake growth relates to faster energy recovery, which was observed to be the case in wind fields with higher turbulence. That is the reason why wakes persist longer in stable atmospheric conditions, where the turbulence is lower (Wu and Porté-Agel, 2012).~~

~~In terms of the turbulence itself, the wake structure is strongly shaped by both the rotor's performance and the atmospheric fluctuations. The former occurs due to blade- and hub-induced vortices, which enhance the turbulence intensity already present due to the atmospheric fluctuations. The vertical turbulence intensity profile is affected by the atmospheric shear, which is why the peak I_{amb} value is present at the top of the rotor. While the smaller-scale eddies are mostly responsible for the energy recovery, the large-scale lateral motions can significantly alter the wake's path, a phenomenon which was termed *wake meandering*. It results in downstream turbines experiencing wake effects intermittently in time, which in turn causes unsteady loads. Experimental studies (España et al., 2011) indicate that meandering seems to occurs only if the incoming flow has eddies much larger than the rotor diameter.~~

1.1 ~~Wind farm flow physics~~

~~The flow entering a cluster of wind turbines undergoes additional transformations compared to the ones discussed in Sect. **??**. For the first several rows of turbines, the flow is highly heterogeneous due to the individual wakes being formed and interacting with each other. This region is referred to as *entrance and flow development region*. While moving further downstream, the flow remains heterogenous at the turbine height, but the flow in higher regions of ABL eventually stabilise, allowing for energy recovery and thus power production to be more balanced. This situation, being only possible for very large wind farms, is referred to as *infinite wind farm case* (Porté-Agel et al., 2020).~~

~~Of prime interest for the project is the first of the above-mentioned flow regions, as the wakes there are not heavily mixed with each other yet and the application of wind farm flow control is much more feasible. The power degradation is strongly pronounced while moving downstream through the first few rows of the wind farm. A study by Barthelmie et al. (2010) presented the impact of wind turbine wakes to the energy production in Horns Rev and Nysted offshore plants, located on the North Sea and Baltic Sea, respectively. The field measurements were done for inflow from different wind directions, analysing the power loss across 10 (Horns Rev) and 8 (Nysted) first turbine rows. The study showed a significant drop in power~~

~~production - the turbines located the 'deepest' in the wind farm would only have the output of approximately 45% or 70% of the first-row-turbine value, depending on the wind direction. Where the inflow angle would cause the turbines in a column to be directly downstream from each other, the major drop would be immediately after the first row, and remain fairly similar. In other cases, the gradually mixing wakes would result in a more gradual drop in power production. This study, along with others (Walker et al., 2016; Nygaard, 2014), proved the great importance of wake distribution to the power production of a wind power plant.~~

1.1 ~~Numerical turbine/farm wake models~~

~~The purpose of engineering wake models, used extensively in both academia and industry, is to estimate the wake development and its effect on the downstream turbines - which then can be used to approximate the power output, machine loads and other key parameters. These approaches rely on implementing necessary simplifications to model the turbine flow with low computational cost. The Jensen model (Jensen, 1983; Katić et al., 1986), despite being developed more than 40 years ago, is still widely used in the industry and provides a basis for more advanced models. It considers a simple top-hat velocity deficit profile based on thrust, rotor radius, downstream distance and a wake expansion coefficient, the latter dependent on aspects such as atmospheric conditions, topography etc. Newer models tend to consider the wake velocity distribution to be a Gaussian profile, which is a decision based on experimental studies of the wake - a study by Bastankhah and Porté-Agel (2014) is the one that gained most attention in that regard. A more sophisticated approach was taken by Ainslie (1988), who derived the model from the approximated Navier-Stokes equations. The model copes well with representing the wake interaction with the ABL, which is due to consideration of turbulent eddy viscosity, analysing how the different turbulent eddies react with each other. In a recent study (Kim et al., 2018), a variant of the Ainslie model proved to be more accurate than the popular Jensen or Bastankah-Porté-Agel models in matching experimental wind tunnel data.~~

~~The Ainslie model was the first to consider both the small and large scales of the turbulence in the context of wake development. It was an inspiration for the more recent Dynamic Wake Meandering (DWM) model (Larsen et al., 2007), which is currently the state-of-the-art for wind turbine load and power modelling, recommended in the latest IEC standard (IEC, 2019). In its essence, the DWM model considers the wake as a passive tracer being moved laterally and vertically by atmospheric motions. It is achieved by turning the wake into a cascade of wake deficits, being 'released' by the turbine with high frequency, moving with an advection speed equal to the ambient flow speed. The path of each tracer is dictated by large-scale stochastic turbulence, in the form of eddies with a characteristic length of twice the rotor diameter or larger. The defined tracer path creates a meandering frame of reference, in which the Ainslie-based velocity deficit and turbine-induced turbulence effects are then separately captured. The result is a model which distinguishes the impact of turbulence: overall wake path dictated by large-scale eddies, and wake energy recovery ordered by small-scale ones. The model requires some form of stochastic turbulence field mimicking the atmospheric conditions, such as Kaimal model (Kaimal et al., 1972) or Mann model (Mann, 1994).~~

1.1 ~~State-of-the-art in wind field reconstruction~~

~~Most of the recently developed inflow estimation methods which aim to capture more dynamic wind phenomena without using external equipment such as LIDARs are based on incorporating the rotor loads. The widely discussed in the literature method introduced by Bottasso and Schreiber (2018) relies on using the out-of-plane blade bending moments to approximate the local inflow wind speed at a given azimuth position. The blades act here as a tool to sample the spatially non-uniform incoming wind field. The rotor is separated into four sectors, and the wind speed values at each are estimated from the measurements using an estimator based on Kalman filtering – this way, the vertical and horizontal shear profiles on the rotor plane can be effectively approximated. The approach was later tested in the field (Schreiber et al., 2020) using a setup of two wind turbines and a met mast, showing an excellent agreement in the shear estimations between the two.~~

~~Another method based on implementing the blade loads, first described by Bottasso and Riboldi (2014), considers using the once-per-revolution (1P) harmonics instead of the time series. The authors proved that the vertical shear and the lateral misalignment in incoming inflow result in a specific ‘trace’ in the frequency domain. The estimator built on this basis was later extended (Cacciola et al., 2016), the outcome being an estimation model that maps the in-plane and out-of-plane 1P blade load harmonics to inflow conditions. In a field application, the model needs to be tuned using the measurements from a met mast and wind turbines in question (Bertelè et al., 2021). The wind field estimates are then produced from the live data using the least squares method or by implementing a Kalman filter. Initial field testing was performed with training data acquired from 10 Hz raw signals averaged into 1 min and 10 min values, showing a remarkable correlation between the ‘sensed’ values and these measured using classic mast-installed sensors. The authors pointed out that for more reliable results, there is a need for a more robust evaluation method than a met mast, which could be potentially solved using a LIDAR.~~

~~The legacy version of the above so-called ’harmonic observer’ proved to have several limitations in its first form, with the most significant being the need for simultaneous measurements of all four ‘wind states’. That would often lead to cross-contamination of data when e.g. one of the sensors is faulty. Another drawback is the way of dealing with nonlinearities between load signals and wind parameters. In a recent study (Kim et al., 2023), the authors decided to overcome these problems by applying neural networks to the wind and loads correlation model. A separate feed-forward neural network was assigned to each of the wind states, thus achieving a separation of the ‘wind sensing’ actions. The performance of the estimator was tested on the same field site as the legacy method (Bertelè et al., 2021), showing an overall similarly high accuracy – the networks trained with field data were able to follow the fast-changing behaviour of the wind field.~~

~~A similar approach to the legacy Bottasso & Riboldi method (Bottasso and Riboldi, 2014) was discussed in (Simley and Pao, 2016) . The authors describe a hub-height wind speed and linear shear components estimator based on the Kalman filter. The turbine data channels used for constructing a state-space model which would be the input for the estimator consisted, among generator rotational azimuth angle and nacelle acceleration, of root bending moments expressed in a non-rotational frame. Tested with a 5 MW turbine numerical model under realistic turbulent inflow conditions, this solution proved to provide an accurate approximate estimation of the above-mentioned wind field properties.~~

~~In another reference (Liu et al., 2021) the authors use the blade loads in a different approach – a Subspace Predictive Repetitive Estimator is proposed to estimate the effective wind speed ‘felt’ by each blade along the span. It is done through the mapping process between the out-of-plane bending moments and wind speed, which can in turn be used to detect wake~~

~~velocity deficits and thus, wake impingement effects. An interesting aspect of this method is the ability to capture partial wakes by measuring the wind speed change along the blade, showing high potential for implementation in inflow wake estimation. That being said, the FLORIS wake model used in the study is an overly simplified representation of an actual wind field, being a simple circular deficit with clear borders and centre; a more realistic study would need to be performed in order to decide the usefulness of this approach in real-life wind farm control.~~

~~This discussion can be summarised with a following statement: the area of wind field reconstruction has undergone a major development over the past years, paving the way for various control solutions that employ the information on the instantaneous and time-average state of the wind flow at the turbine. One of such applications is estimation of the wake impingement conditions, which is discussed in the next section.~~

1.1 ~~State-of-the-art in wake detection and characterisation~~

~~So far, the methods that attempted to obtain detailed information on incoming flow, specifically focused on detecting or characterising wakes, were divided into two categories: LIDAR-based and load-based.~~

~~Light Detection and Ranging (LIDAR) devices can provide an accurate representation of the wind farm flow due to their ability to scan the wind field at high frequency. Some examples describing their use for wake properties analysis include the works by Bingöl et al. (2010), Trujillo et al. (2011) and Conti et al. (2020). In these studies, the authors managed to successfully capture such effects as velocity deficit, added turbulence and meandering, proving the usefulness of this technology for estimating detailed wake properties. For this reason, the application of a nacelle-mounted LIDAR in dynamic wake characterisation was recently investigated, using both an upwind-facing device (Lio et al., 2020) as well as with a downwind-facing setup (Raach et al., 2016). In both these studies, the measurements were used in combination with a numerical wake model, ultimately giving an instantaneous wake center position estimate which in turn could be used for closed-loop flow control.~~

~~A recent branch of research tackles wake detection and characterisation using information obtained purely from the turbine operational data. The basic idea behind this method is that a variation in the incoming wind field can be identified with changes in the turbine response, such as blade root bending moments and generator speed. The previously discussed in Sect. **??** approach of dividing the rotor into four quadrants and estimating the respective wind speeds was specifically tested for wake characterisation (Bottasso and Schreiber, 2018; Schreiber et al., 2016) with the idea of using the local wind speed estimates to continuously optimise an engineering wake model. Its performance was shown to be limited by several problems, most potentially due to the rather simplistic nature of the estimation in this method. The works of Onnen et al. (2022) and Dong et al. (2021) are great examples for employing the out-of-plane blade loads in a model based on the extended Kalman filter estimation, which effectively allows for dynamic wake center position tracking in both lateral and vertical directions. In another study by Farrell et al. (2022), the authors implemented a recurrent neural network trained with experimental and simulation data to estimate the lateral position of the wake centre. These models have proved to offer high dynamic wake characterisation accuracy comparable to LIDAR-based methods, while remaining highly sensitive to higher turbulence intensity values, as well as varying veer and shear conditions.~~

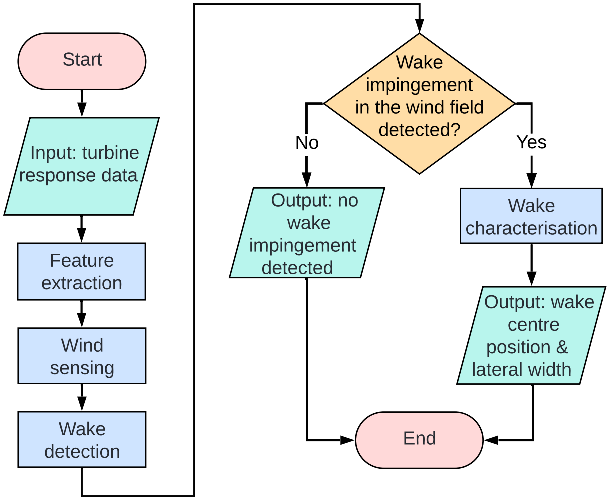


Figure 1. A flowchart describing the proposed wake estimation framework.

~~Having presented the state-of-the-art in the field, we now restate the research gap: the majority of up-to-date approaches focused primarily on wake characterisation, with a key assumption that a turbine is continuously impinged by a wake deficit. The central part of the generalised method we propose in this work is a wake detector aiming to 'sieve out' the wake-impinged flow from the incoming wind field, and thus applying wake characterisation in a much better informed manner.~~

2 Methodology

Section 2.1 discusses the layout of the developed framework and its high-level characteristics. The details of the implementation are described in Sect. 2.2 to 2.5. ~~Lastly, the specific methods used for testing the wake sensing performance are described in Sect. ??.~~

2.1 A summary of the proposed framework

Figure 1 ~~presents the~~ shows the information flow between the models within the proposed wake estimation framework~~in the form of a flowchart. The framework consists of several processes taking part in implementing the input turbine response data towards obtaining the ultimate information on the estimated wake impingement conditions.~~. Figure 2 presents two examples that illustrate the framework's performance when (a) the wind field doesn't have a clear wake impingement from a nearby turbine, (b) when it does. The high-level details of the individual models that the framework consists of are discussed below.

The first constituent model in the developed framework is a wind field estimator, which when provided with turbine response time series as the input, is capable of producing the wind field representation, thus showing the flow at the rotor as the output. This process is hereafter referred to as *wind sensing*. A successfully generated wind field estimation provides suf-

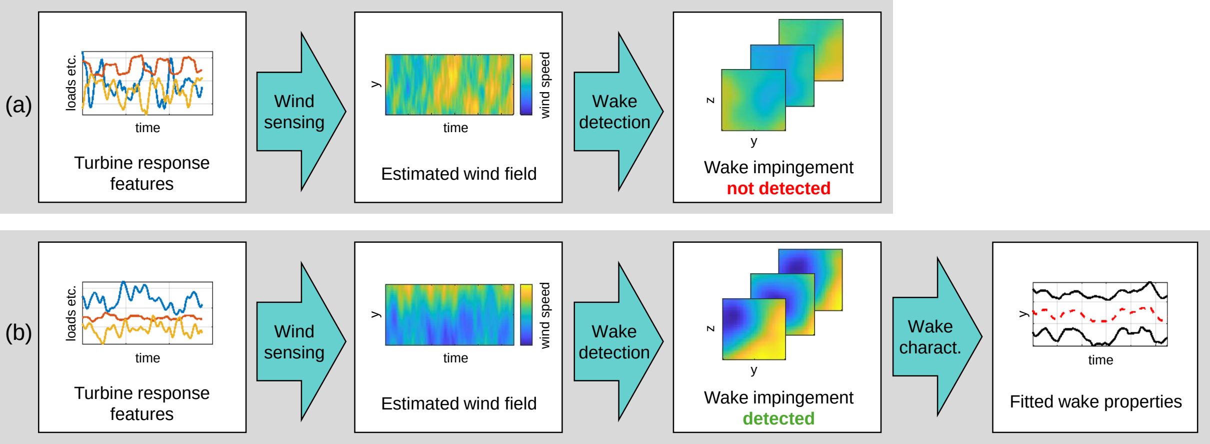


Figure 2. Two examples showing the framework's performance; (a) the incoming wind field doesn't have a clear wake deficit, thus the framework doesn't go into the wake characterisation; (b) there is a clear wake deficit, thus the final step of wake characterisation is implemented.

ficient information to analyze the flow field and extract the desired wake impingement information. For this framework, we propose to perform wake detection using a ~~convolutional neural network~~ Convolutional Neural Network (CNN), which, after being trained, can distinguish the flow with a well-defined wake deficit from a nearby turbine. If a likely wake impingement is detected, a suitable algorithm aiming to estimate the wake's properties is to be implemented. The ultimate framework output data are two time series of the wake centre position and wake lateral width, describing the horizontal distance where the degree of wake impingement is significant. The combination of these two time series ~~is a simplified , but qualitatively sufficient 2D~~ provides a simplified two-dimensional representation of a meandering wake deficit. ~~The developed methodology also allows to extract a time series of peak wake deficit ΔU value, however this is is not discussed in this work. This representation could be used~~ This representation can serve as an input to ~~the~~ wind farm flow control techniques ~~which~~ that are informed by the lateral properties of the impinging wake, such as ~~the ones using yaw as the~~ those that use yaw as a control action (Howland et al., 2020; SGRE, 2019). Focusing specifically on the lateral properties is also motivated by the fact that the wake meandering is more pronounced in the transverse direction than in the vertical direction, with the ratio reflecting the turbulence intensities of the respective components (Yang and Sotiropoulos, 2019).

2.2 ~~Framework implementation: training data acquisition~~

~~A dataset comprised of 1200 simulations was produced, each one constituting of generating a unique wind field and calculating the corresponding~~ The developed methodology also allows to extract a time series of ~~instantaneous turbine response. Synthetic wind fields were acquired by firstly generating ambient turbulence boxes using the Mann spectral tensor model (Mann, 1994) and then superimposing them with wakes generated with the SGRE's in-house implementation of the DWM model (Larsen et al., 2007) . This version of the model follows the parametrisation suggested in the latest IEC standard (IEC, 2019). The turbine operation~~

~~for the simulated conditions was then calculated using the aeroelastic code BHawC.~~ peak wake deficit ΔU value, however this is is not discussed in this work.

~~The simulations were performed for a virtual wind farm located in the North Sea. In order to train the wake detection model, turbine location and wind direction combinations were chosen such that test cases clearly differentiated between the effects of wake impingement and standalone atmospheric turbulence. With that in mind, two neighbouring turbines of interest, located near the northern edge of the wind farm , were selected~~

2.2 Framework implementation: training data acquisition

The simulation environment enables modeling the operation of a given turbine at any location within the wind farm layout, capturing wake impingement effects from multiple turbines simultaneously. While this aspect plays a key role during testing, training is limited to interactions between two devices located on the farm border: a first-row ~~turbine ’emitting’ the wake~~ wake ’emitter’ and a second-row ~~turbine ’receiving’the wake. It was crucial for these turbines to be at the edge of a wind farm to ensure that only one wake is impacting~~ wake ’receiver’, spaced approx. 6 rotor diameters apart. The role of the ’receiver’ ~~turbine; with more wakes having a chance to superimpose with each other, the impingement effects would be more complex and the trainingof a wake detection model would become more ambiguous~~is to provide the turbine response data for training. Four wind directions, representing four distinctive wake impingement scenarios ~~, were~~ as seen from the front of the rotor, are defined as follows: a) fully impinged, turbine directly downstream from the closest neighbour; b) partially impinged, wake impacting the left side of the rotor; c) partially impinged, wake impacting the right side of the rotor; d) no impingement. The wind direction differs by 5 degrees between the fully and partially impinged cases. This setup allows to clearly differentiate between the effects of full/partial wake impingement and standalone atmospheric turbulence. It ensures that only one wake is impacting the ’receiver’ turbine~~would provide the turbine response data for training. The example wind farm layout as well as~~; including multiple wakes would complicate the impingement effects and make it more challenging to define wake detection classes (see Sec. 2.4). The partial layout of the wind farm and the wind directions defining the four classes ~~have been~~ are shown in Fig. 3.

~~The ambient inflow wind speed U_{amb} was set as a range of 100 equally distributed values between 5 m/s and 25 m/s;~~ A dataset comprised of 1200 simulations is produced, with simulation parameters and run counts presented in Table 1. Ambient wind speed U_{amb} and ambient turbulence intensity ~~I_{amb} was set at three values of 3%, 5% and 7%. This was done~~ I_{amb} values are varied to ensure that the simulated conditions would represent a multitude of different offshore wind states. ~~Due to the fact that ambient wind fields were acquired~~ Each simulation consists of generating a unique wind field and calculating the corresponding time series of instantaneous turbine response. Synthetic wind fields are acquired by generating ambient turbulence boxes using the Mann ~~turbulence~~ spectral tensor model (Mann, 1994) and superimposing them with wakes generated with the SGRE’s in-house implementation of the DWM model (Larsen et al., 2007). The version of the model used follows the parametrisation suggested in the latest IEC standard (IEC, 2019). In its essence, the DWM model considers the wake as a passive tracer being moved laterally and vertically by stochastic turbulence with a characteristic length of twice the rotor diameter or larger. The turbine operation for the simulated conditions is calculated using the aeroelastic code BHawC. Since ambient wind fields

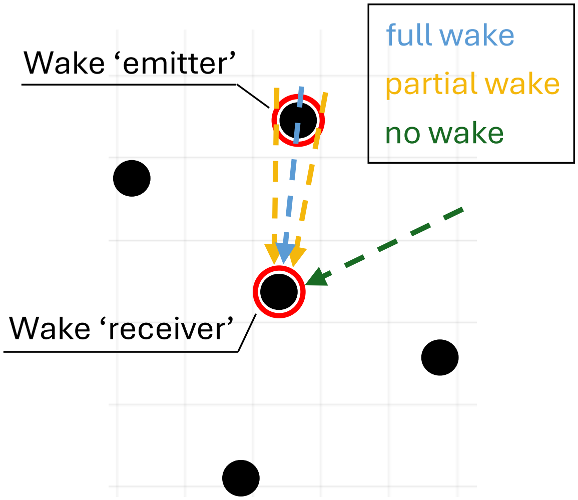


Figure 3. Simulated wind farm with highlighted locations of two turbines of interest and selected wind directions for the training process.

Table 1. Training simulation subsets. For each wind direction, 300 simulations capturing all U_{amb} and I_{amb} combinations are performed.

Wind direction [°]	Impingement case	U_{amb} [m/s]	I_{amb} [%]	# of simulations
9	Full	5-25 (step size 0.2)	3, 5, 7	300
4	Partial left			
14	Partial right			
45	None			
				Total: 1200

are acquired with the Mann turbulence model, the atmospheric conditions ~~were~~ are assumed to be neutral (Mann, 1994). The presence of vertical wind shear ~~was~~ is simulated using the wind profile power law, with the shear exponent value of 0.07 selected for all cases. ~~It should be also mentioned that the turbines were~~ Turbines are simulated with no yaw angle misalignment.

2.3 Framework implementation: wind sensing

2.3.1 Wind sensing input and output data selection

Figure 4 illustrates ~~the process for training~~ training of the wind sensing model and demonstrates its ~~post-training performance in producing wind field estimations~~application scheme. By providing a varied dataset of turbine response time series and the

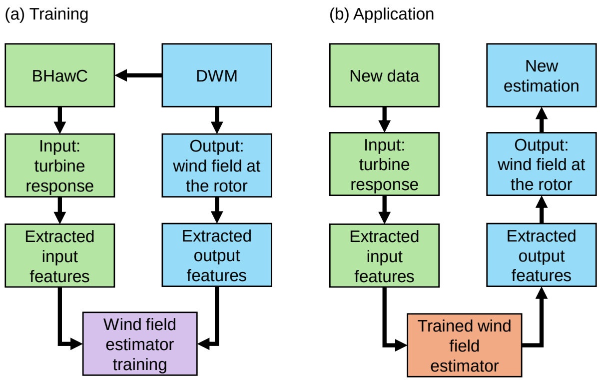


Figure 4. Simplified diagram visualising the flow of data during (a) training, (b) application of the wind sensing model.

corresponding wind field representations, the estimator ’learns’ to approximate the wind speed distribution at the rotor from the turbine sensor data.

The ~~capabilities of the aeroelastic code allow to get~~ aeroelastic code provides access to an extensive set of turbine response signals~~like acceleration of different turbine components, generator parameters etc. To~~ , such as component accelerations or generator speed. A conditional dependence analysis (readers further interested in the specific method used are referred to (Azadkia and Chatterjee, 2021)) is undertaken to determine which data channels are best coupled with the variations in the flow~~, a conditional dependence was undertaken. Readers interested in the specific method are referred to (?). As a result, following turbine response data was used as the wind sensing inputs~~. The algorithm tests different combinations of turbine signals, estimating the likelihood that wind speed can be deterministically predicted from them. The following combination with the strongest overall correlation to wind speed is identified and used in wind sensing: blade root bending moments in both the flap-wise and edge-wise direction; ~~pitch angles; generator rotational speed (rpm); azimuth angle (used only~~ mean pitch across all three blades, simply calculated as $\overline{\beta} = (\beta_1 + \beta_2 + \beta_3)/3$; rotor rotational speed ω. Rotor azimuth angle and individual pitch angles are also extracted to use for transformations of other inputs~~).~~ .

The wind field is considered to be a three-dimensional Cartesian grid, with three axes: X representing streamwise distance, Y representing lateral distance, and Z representing vertical distance. To represent the wind speed distribution in the flow, each point on that grid stores the local values of three wind speed components U, V and W that correspond to the spatial variation in the X, Y and Z ~~axis~~axes, respectively. For this specific application, the ~~approach taken was to consider the wind field as~~ wind field is represented with the spatial distribution of the longitudinal wind speed component U, as it is the one that is primarily impacted by wake deficit (Dimitrov et al., 2017). ~~As a result, the wind fields used in the wind sensing are essentially~~ All models treat the wind field as a time series of YZ slices showing the chronological sequence of ~~U~~ U distributions at the rotor plane.

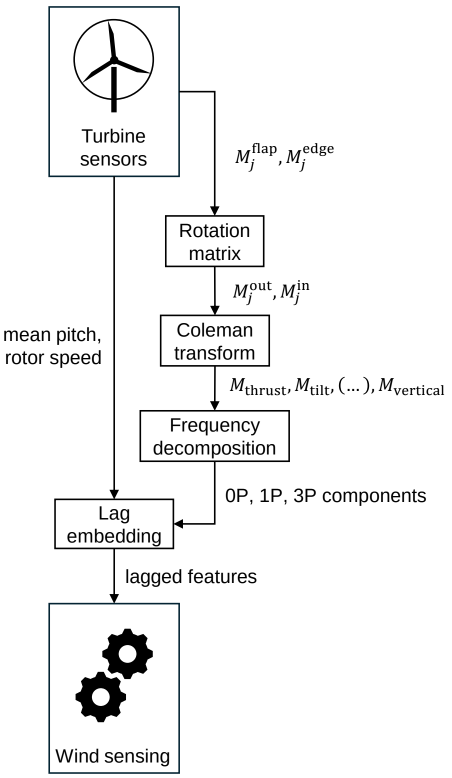


Figure 5. Block diagram of the turbine response processing. $M_j^{\text{flap/edge}}$: flap/edge blade root bending moments, $M_j^{\text{out/in}}$: in/out of plane blade root bending moments, j: blade index, $M_{\text{thrust/tilt/(...)/vertical}}$: rotor loads. Detailed description in the text.

2.3.2 Wind sensing input processing

Selected turbine response time series ~~(blade root bending moments, mean pitch and rpm) were then~~ are processed to extract the optimal features from the raw data. The processes described in this section ~~were~~ are applied both during training and ~~the~~ performance testing - they are an integral part of the data 'pipeline'. ~~First of all~~Figure 5 illustrates the transformations taking place in the feature extraction.

First, using the pitch signals, the edge- and flap-wise blade root bending moments ~~were~~ are transformed to in- and out-of-plane coordinates. Performing the rotation between the two frames ~~was~~ is calculated for a given blade j as follows:

$$\begin{bmatrix} M_j^{\text{out}} \\ M_j^{\text{in}} \end{bmatrix} = \begin{bmatrix} \cos(\beta_j) & \sin(\beta_j) \\ -\sin(\beta_j) & \cos(\beta_j) \end{bmatrix} \begin{bmatrix} M_j^{\text{flap}} \\ M_j^{\text{edge}} \end{bmatrix} \tag{1}$$

where M_j^{out}, M_j^{in} are out-of- and in-plane root bending moments, respectively; M_j^{flap}, M_j^{edge} are flap and edge root bending moments, respectively; β_j ~~stands for~~ is the instantaneous pitch angle~~for a given blade indexed j~~.

Next, using the azimuth angle time series, the Coleman transformation (Coleman, 1943) ~~was~~ is applied to the blade load data. It ~~allowed~~ allows to effectively map the separate, rotary-frame bending moments from each blade onto a stationary frame of reference. As a result, the dynamic response to the variation in the wind can be expressed with six variables describing the loads for an entire rotor: ~~$M_{thrust}, M_{tilt}, M_{yaw}$~~ $M_{\text{thrust}}, M_{\text{tilt}}, M_{\text{yaw}}$ (for out-of-rotor-plane movement), ~~$M_{torque}, M_{lateral}, M_{vertical}$~~ $M_{\text{torque}}, M_{\text{lateral}}, M_{\text{vertical}}$ (for in-rotor-plane movement). These are ~~henceforth~~ referred to as *rotor loads*. The Coleman transformation is defined as follows:

$$C(\psi) = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 2/3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ \cos(\psi) & \cos(\psi + \frac{2\pi}{3}) & \cos(\psi + \frac{4\pi}{3}) \\ \sin(\psi) & \sin(\psi + \frac{2\pi}{3}) & \sin(\psi + \frac{4\pi}{3}) \end{bmatrix} \tag{2}$$

where ψ is ~~instantaneous~~ rotor azimuth angle. It is used to transform the blade loads as follows:

$$\begin{bmatrix} M_{\text{thrust}} \\ M_{\text{tilt}} \\ M_{\text{yaw}} \end{bmatrix} = C(\psi) * \begin{bmatrix} M_1^{\text{out}} \\ M_2^{\text{out}} \\ M_3^{\text{out}} \end{bmatrix} \tag{3}$$

$$\begin{bmatrix} M_{\text{torque}} \\ M_{\text{lateral}} \\ M_{\text{vertical}} \end{bmatrix} = C(\psi) * \begin{bmatrix} M_1^{\text{in}} \\ M_2^{\text{in}} \\ M_3^{\text{in}} \end{bmatrix} \tag{4}$$

where ~~$M_{1-3,out}$~~ M_{1-3}^{out} is out-of-plane bending moment, and ~~$M_{1-3,in}$~~ M_{1-3}^{in} is in-plane bending moment for blades indexed 1-3.

The rotor load time series are ~~then~~ decomposed into its key frequency components. This process is visualised on Fig. 6. The five frequency components selected via a conditional dependence ~~(?) study are $cos(0P)$~~ (Azadkia and Chatterjee, 2021) study are cos(0P) for describing the mean signal value, ~~$cos(1P)$ and $sin(1P)$~~ cos(1P) and sin(1P) for describing the once-per-revolution variation, and finally ~~$cos(3P)$ and $sin(3P)$~~ cos(3P) and sin(3P) for describing the thrice-per-revolution variation. A sliding window function with a width of three revolutions ~~was~~ is applied, projecting the original load time series onto each of the ~~the~~ frequency components. The Fourier coefficients $\underline{a_1}$ - $\underline{a_5}$ vectors obtained from these projections are ~~then~~ stored, producing a coefficient time-series for each ~~component which eventually become the data used by the wind sensing model~~rotor load.

In order to capture the short-term temporal dependencies and patterns in time series of all wind sensing inputs~~(including pitch and rpm)~~, the features ~~were~~ are embedded with their lagged values. For each sample in a time series, two additional features expressing the past value of the curve ~~were~~ are added. These lagged features ~~were~~ are obtained by shifting the time series by 4 and 8 seconds from the current time stamp. ~~These specific lag values were~~ This effectively makes the estimation more stable and noise-resistant, as the wind slice is reconstructed with turbine response across several seconds. Specific lag values used are determined by testing the framework's performance with a few different configurations and choosing the one that ~~produced~~ gives the best overall results.

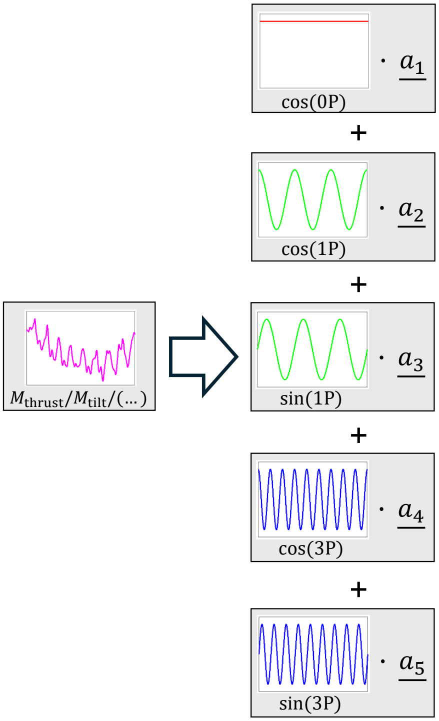


Figure 6. An example of rotor load time-series ~~decompositions~~ decomposition into frequency component coefficient vectors ($\underline{a}_1$ - $\underline{a}_5$) capturing the contributions of each fourier mode at each point in time.

Table 2. A summary of features extracted from turbine response.

Original turbine response signal	Preprocessing methods	Resultant # of features
M^{out}_{1-3}, M^{in}_{1-3}	Rotation matrix, Coleman transform, frequency decomposition, lag embedding	90
$\overline{\beta}$	Lag embedding	3
ω	Lag embedding	3
		Total: 96

Due to implementing the transformations discussed in this section, using the turbine response data as wind sensing input effectively stops being a time series modelling task; instead, it becomes a time-independent regression task. Considering the entire feature extraction process, at each point in time, a 96-dimensional input vector captures and encodes the turbine's current operating state. Table 2 serves as a summary of these features.

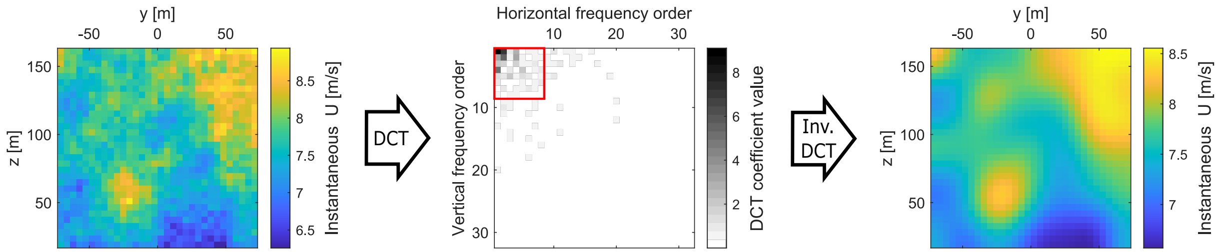


Figure 7. An example showing how the YZ slice of wind field (left) is transformed into a matrix of DCT components (centre) and then back into a truncated slice of wind through an inverse DCT (right). The red square indicates which DCT coefficients are stored for the implemented wind field representation.

2.3.3 Wind sensing output processing

The three-dimensional wind field representations are ~~being processed for the sake of~~ processed for data compression and efficient extraction of ~~its~~ their key features for the best application in wind sensing. ~~As mentioned before, the~~ The wind field is represented with a sequence of YZ slices showing the U distribution at each point along the X axis. The spatial features of each slice are ~~then~~ extracted, capturing the key characteristics of the wind field in a more compact form. The feature extraction is performed using the two-dimensional discrete cosine transform (DCT) (Ahmed et al., 1974). The result is similar to applying the principal component analysis; however, as opposed to this far more complex method, the DCT is a fixed transform and doesn't require any training. The DCT effectively converts the wind slice into a collection of cosine functions oscillating at different frequencies in both the vertical and horizontal directions. Each DCT coefficient corresponds to a specific frequency in the transformed domain and describes the contribution of each frequency in the distribution of U on the given YZ slice. This concept, applied to an example wind field YZ slice, is presented on Fig. 7.

The lower-order frequencies primarily capture the large-scale structures of the field, while the higher-order are more associated with encoding smaller-scale, random fluctuations. It can be seen that the most significant values can be found in the upper left corner of the diagram, meaning the majority of flow distribution is being kept by the lower orders of DCT coefficients. For this reason, the key information on the flow can be easily extracted by storing only the lower orders and disregarding the higher order coefficients. Truncating the collection of DCT order coefficients essentially acts as a spatial low-pass filter. In this work, retention of 8 DCT coefficients ~~was~~ is found to enhance computational efficiencies while allowing for good wind sensing accuracy. The right plot on Fig. 7 shows ~~on an example~~ the effect of truncating the higher frequencies; it can be noticed how the wind slice is significantly 'smoothed', leaving only the major U fluctuations.

2.3.4 Training the wind sensing estimator

Having extracted the features from both the inputs and outputs, two thirds of data ~~was then~~ is used for the wind field estimator training, while the remaining 1/3 ~~was~~ is put aside for generating wind field estimations for wake detector testing. The current implementation utilises a localised linear regression approach ~~, where a collection of simple modelsis fitted to subsets of~~

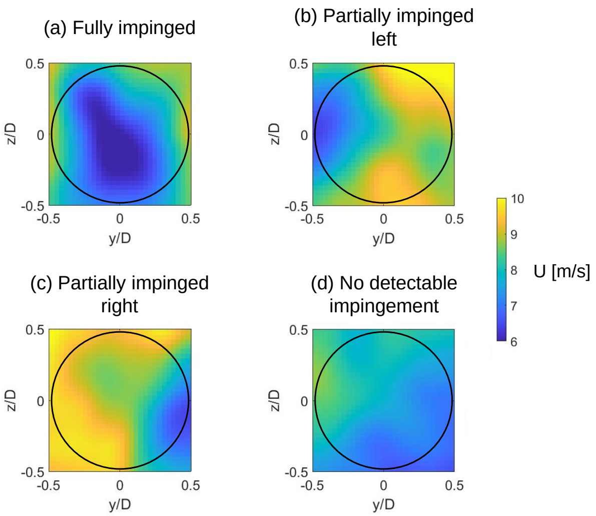
~~input data, building~~ (Cleveland et al., 1988), where the estimator training is performed for multiple localised models. The preprocessed turbine response data is a time series of multidimensional input points, with each dimension representing a different feature (see Table 2). Two of these dimensions, namely unlagged $\overline{\beta}$ and unlagged ω (these being the key variables indicating the turbine's operating point), are used to define a projection plane, allowing input points to be mapped and binned by their values. A set of evenly spaced points on this plane is selected, and a local linear regression model is trained at each. Training data is assigned to models based on proximity within a defined radius, with overlapping points contributing to multiple estimators. Predictions for new inputs are made using the model nearest to their location on the projection plane. As a result, the training process builds functions that correlate the ~~wind sensing~~ inputs to outputs only at a narrow section of data distribution. This ~~allowed to avoid fitting of~~ eliminates the need to fit a global non-linear function that would singlehandedly need to describe the complex relationship between wind and turbine signals. ~~The data points are assigned to training of a specific local linear regression model based on their pitch and rpm value, these being the key variables indicating the turbine's operating point.~~

2.4 Framework implementation: wake detection

Considering that wake detection can be essentially brought down to a classification task, we propose to deal with it by implementing a ~~Convolutional Neural Network~~convolutional neural network. The overview of this deep learning technique can be found in Appendix A. The approach taken ~~was~~ is to operate on the YZ snapshot samples (so-called 'front view'), showing the estimated instantaneous wind component U across the whole rotor plane, which ~~made~~ makes it possible to easily differentiate between impinged and non-impinged conditions. To capture the different forms of wake impingement accurately, a total of four classes~~were defined~~, directly relating to four wind directions used during the training data acquisition (see Sect. 2.2), are defined. These are as follows: a) fully impinged, turbine directly downstream from the closest neighbour; b) partially impinged left, wake impacting the left side of the rotor; c) partially impinged right, wake impacting the right side of the rotor; d) no detectable impingement. Figure 8 shows selected YZ samples being appropriate examples of each one of the four classes.

The turbine response time series from the remaining 400 simulations, set aside during the wind sensing estimator training, ~~were~~ are now used for generating wind field estimations, allowing to build a new dataset for the CNN training. An extensive manual review of these wind fields ~~allowed~~ allows for a following conclusion: in the available dataset, a clear, non-dissipated wake is only visible for lower values of the selected ambient wind speed range - typically up to 14-16 m/s, depending on the ambient turbulence intensity. The aim of the detector is to recognize (as robustly as possible) whether the turbine is most likely experiencing a wake impingement from a nearby machine, which is why ~~it is crucial that~~ the training data ~~is~~ must be devoid of samples that would be arguable regarding which class they should be part of. With that in mind, the CNN ~~was~~ is trained with data from simulations where the mean ambient wind speed ~~U_{amb}~~ U_{amb} is between 5 and 15 m/s, while maintaining the full range of turbulence intensity values (3%, 5% and 7%). The YZ 'snapshot' samples ~~were~~ are taken every 10 seconds from the available training data, resulting in 11,200 labelled samples.

~~The training was conducted using Matlab's Deep Learning Toolbox. The~~ The implemented CNN architecture ~~was based on~~ is presented in Table 3. The input layer has the same resolution as the YZ slice of estimated wind. The three convolution layers ~~and two max pooling layers~~have iteratively increasing number of kernels, allowing the network to progressively learn


(a) Fully impinged

(b) Partially impinged left

(c) Partially impinged right

(d) No detectable impingement

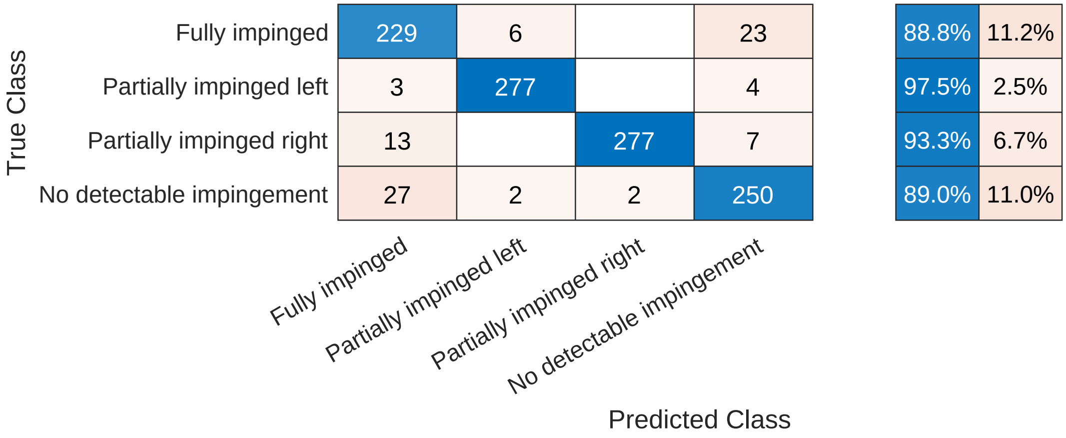
Figure 8. Example front-view YZ snapshot samples representing four classes defined for the neural network training. The rotor area indicated with a black circle.

more complex features at deeper levels (Goodfellow et al., 2016). A batch normalisation layer ~~was~~ is added after each of the convolution layers for the sake of normalisation of weight gradients and neuron activations, thus speeding up the training process. Two max pooling layers reduce the spatial size of the feature maps. The activation function ~~was~~ is selected as the Rectified Linear Unit (ReLU). The final three layers transform the learned features into predictions, connecting all of the neurons, normalising outputs into probabilities and performing final classification. An extended description of the layers used can be found in Appendix A. The optimal hyperparameter values are chosen based on trial-and-error, monitoring the CNN accuracy through integrated testing.

The training is conducted using Matlab's Deep Learning Toolbox. 90% of the available labelled data (10,080 samples) ~~was~~ is used for training. The samples ~~were~~ are shuffled before the training process, as well as after each one of the four epochs. The training ~~was~~ is repeated several times to make sure there are no deviations. The remaining ~~1,120 samples were~~ 10% of samples are used for an integrated testing cycle, which automatically ~~occurred~~ occurs after the training ~~was complete. The average accuracy of the trained network in recognizing a wake impingement case turned out to be approximately 91%~~is complete. Figure 9 shows its results in the form a confusion matrix. As can be observed, the main source of error are the misclassifications between the 'Fully impinged' and 'No detectable impingement' classes. This phenomenon is further investigated in Sect. 3 and 4, where the performance of each of the models is analyzed in more detail under a variety of conditions.

Table 3. Implemented CNN architecture.

Layer #	Name	Details
1	Image input	Layer size = 32x32
2	Convolution	8 kernels with size = 3
3	Batch normalisation	
4	ReLU	
5	Max pooling	Pool size = 2, stride = 2
6	Convolution	16 kernels with size = 3
7	Batch normalisation	
8	ReLU	
9	Max pooling	Pool size = 2, stride = 2
10	Convolution	32 kernels with size = 3
11	Batch normalisation	
12	Fully connected	
13	Output normalisation	Type: softmax
14	Classification	Layer size = 4



True Class \ Predicted Class	Fully impinged	Partially impinged left	Partially impinged right	No detectable impingement		
Fully impinged	229	6		23	88.8%	11.2%
Partially impinged left	3	277		4	97.5%	2.5%
Partially impinged right	13		277	7	93.3%	6.7%
No detectable impingement	27	2	2	250	89.0%	11.0%

Figure 9. Confusion matrix for the integrated testing of trained CNN. Each row represents the instances of a true class, whereas each column represents the instances of a predicted class. As a result, the values in diagonal (blue) and off-diagonal (orange) cells correspond to correct and false classifications, respectively. Percentages of correct (blue) and false (orange) classifications within each true class shown on the right.

2.5 Framework implementation: ~~Wake~~ wake characterisation

The final part of the developed framework is the identification of the key wake parameters in cases where a clear wake impingement is detected. This ~~was~~ is undertaken by scanning for the characteristic velocity deficit shape in the appropriate wind field representations.

2.5.1 ~~2D~~ Two-dimensional Gaussian fitting

The 'tracking' of wake position and width ~~was~~ is achieved by implementing a ~~2D~~ two-dimensional Gaussian fitting scheme based on a least squares algorithm. This method is a commonly used solution in the literature for the dynamic wake properties analysis (Abkar and Porté-Agel, 2015; Trujillo et al., 2011; Conti et al., 2020). ~~Analogically to wake detection, the fitting algorithm was applied to a time series of YZ slices representing the U distribution at the rotor plane, performing the non-linear least square fit based on functionalities from Matlab's Optimization Toolbox~~First, the wake deficit is defined for every sample as follows:

$$U_{\mathrm{def}}(y,z) = U(y,z) - U_{\mathrm{amb}} \tag{5}$$

where $U(y,z)$ is the instantaneous wind speed distribution across the YZ plane, and U_{amb} is the mean ambient wind speed value.

A general, bivariate ~~2D~~ two-dimensional Gaussian function for spatial variables of ~~y_i and z_i~~ y and z is expressed as:

$$f_{\underline{G,bivariate}\,\text{bivariate}}(y_{\underline{i}}, z_{\underline{i}}) = \frac{A}{2\pi\sigma_y\sigma_z\sqrt{1-\rho^2}} \exp\Bigg[-\frac{1}{2(1-\rho^2)}\Bigg(\underline{\frac{(y_i-y_c)^2}{\sigma_y^2}}\,\frac{(y-y_c)^2}{\sigma_y^2} - \underline{\frac{2\rho(y_i-y_c)(z_i-z_c)}{\sigma_y\sigma_z}}\,\frac{2\rho(y-y_c)(z-z_c)}{\sigma_y\sigma_z} + (z_i \tag{6}$$

where A is amplitude equal to height of the peak; σ_y and σ_z are standard variations along Y and Z axis, respectively; ρ is correlation coefficient between function's spread in Y and Z; y_c and z_c are means along Y and Z, respectively.

~~In the context of fitting the function to the velocity wake deficit, respective Gaussian parameters can be used to acquire~~ The time series of wake deficits from Eq. (5) is fitted to Eq. (6) through a Levenberg-Marquadt non-linear least squares algorithm implemented with Matlab's Optimization Toolbox. Fitted Gaussian parameters are then used to identify wake properties: the function peak at (y_c, z_c) can be interpreted as position of wake centre; σ_y and σ_z describe the spread of wake along Y and Z axis, respectively; ρ can be used to calculate the ~~rotation angle as well as the~~ lengths of semi-minor and semi-major axis of the wake ellipse~~. A Levenberg-Marquadt non-linear least-squares algorithm was employed to optimize the parameter set ($A, \sigma_y, \sigma_z, \rho, y_c, z_c$) towards a best fit with a snapshot data. The initial guess for a function peak at (y_c, z_c) was defined as the minimum U value location~~, as well as their orientation with respect to the YZ axes. The initial guess for ~~standard deviations~~ σ_y and σ_z ~~was defined as 0.5D~~ is defined as half of rotor diameter D in order to best reflect the typical width expected of a wake.

As explained earlier, the framework outputs refer solely to the lateral wake characteristics, which are expressed by the variation on the Y axis. Out of all fitted parameters, only two ~~then~~ are required to describe the ~~wake variation in the lateral~~

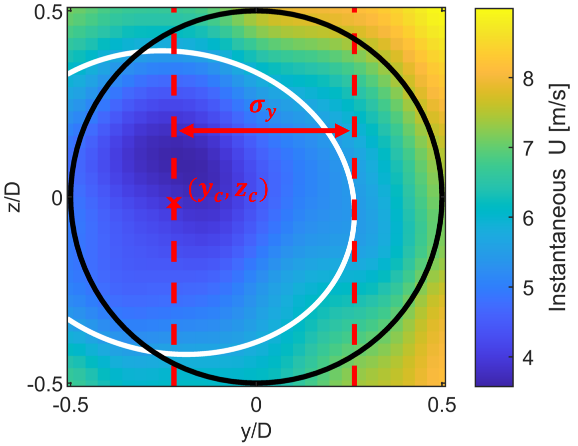


Figure 10. An example of a bivariate ~~2D~~ two-dimensional Gaussian fit performance applied to a YZ wind field sample. Rotor outline marked with black line; fitted wake ellipse marked with white line; fitted wake centre at (y_c, z_c) marked with red x sign; the extent of fitted lateral half-width σ_y marked with dashed lines.

~~direction~~lateral wake variation: y_c and σ_y. To ~~express~~ parametrically represent the size of the wake deficit~~parametrically, an assumption is made~~ , it is assumed that the lateral half-width ~~is equal~~ corresponds to standard deviation in the Y direction. With that in mind, the wake extends laterally for $2\sigma_y$, with the fitted centre y_c in the middle. Figure 10 shows an example of the bivariate ~~2D~~ two-dimensional Gaussian fit applied on an YZ wind field sample, with the key characterised properties indicated.

2.5.2 Filtering with moving average window

The lateral wake half-width and centre position time series ~~were filtered~~ are smoothed using a moving average ~~filtering scheme to smooth fluctuations originating from combining the independent results from each snapshot. It was done in way that would distinguish the slower variation being a result of wake meanderingand attenuate the~~ filter to reduce high-frequency ~~peaks~~noise from sample-by-sample variability, while preserving slower variations associated with wake meandering. This way, the ultimate characterised wake properties time series ~~y_c^{filt} and σ_c^{filt} were~~ y_c^{filt} and σ_y^{filt} are defined as follows:

$$\begin{bmatrix} y_c^{\text{filt}}(t) \\ \sigma_y^{\text{filt}}(t) \end{bmatrix} = \frac{1}{\tau} \int\limits_{t-\tau}^{t} \begin{bmatrix} y_c(t) \\ \sigma_y(t) \end{bmatrix} dt \tag{7}$$

The variable τ stands for window size in seconds which is calculated separately for every wind field in a following manner:

$$\tau = \frac{L_{meand}}{U_{hh}} \frac{L_{\text{meand}}}{U_{\text{hh}}} \tag{8}$$

where ~~L_{meand}~~ L_{meand} is the characteristic meandering length equal to two rotor diameters (value corresponding to findings from the literature, used in several wake models like the DWM itself (Larsen et al., 2007)) and ~~$\overline{U_{hh}}$~~ $\overline{U_{\text{hh}}}$ is the mean U value at the hub height.

It should be noted that in this method of moving average filtering, the averaging window is not centred at the current sample; instead, the window extends from the past sample occurring τ seconds ago to the current sample. As a result, the filtered wake centre and lateral half-width time series experience a small lag of $\tau/2$.

2.5.3 Treatment of non-impinged samples

To satisfy the assumption that wake characterisation should not be performed for wind samples without clear impingement, the filtered wake properties are treated for a given time stamp i as follows:

$$y_c^{\text{filt}}(i), \sigma_y^{\text{filt}}(i) = \begin{cases} y_c^{\text{filt}}(i), \sigma_y^{\text{filt}}(i), & \text{if class}(i) = \text{fully/partially impinged} \\ \text{NaN}, & \text{if class}(i) = \text{no detectable impingement} \end{cases} \tag{9}$$

The assignment of NaN values ensures that the samples identified as non-impinged are not being represented with a two-dimensional Gaussian wake in the framework output.

2.6 ~~Performance testing~~

~~For the purpose of testing the performance of the developed framework , several sets of simulations showing its performance for all wind directions were conducted. Each simulation set generated 360 unique wind fields, each~~

3 Results

3.1 How is the performance evaluated

The framework's performance is evaluated with 3600 new simulations divided into 10 subsets. The configurations with number of runs is shown in Table 4. Each simulation is paired with the corresponding turbine response~~, one for each wind direction angle. The U_{amb} and I_{amb} values were varied across the simulation sets, with ranges of 5, 7, 9, 11, 13, 15 m/s for the former and 3%, 5%, 7% and 9% for the latter. All simulations were~~. All simulations are performed for one 'receiver' turbine, which ~~was~~ is more centrally located within the wind farm compared to the machine used for the training; by doing so, it experiences different types of impingement for different wind directions. ~~Taking all of these into consideration, the performance testing dataset consisted of 3600 simulations. It was used to undertake the performance analysis of~~ Both qualitative and quantitative performance analysis is conducted for each of the three constituent models~~, the results of which are discussed in Sect. 3.~~.

4 ~~Results~~

Table 4. Testing simulation subsets. For each U_{amb} and I_{amb} combination, 360 simulations capturing full wind direction spectrum are performed.

U_{amb} [m/s]	I_{amb} [%]	Wind direction [°]	# of simulations
5, 7, 9, 11, 13, 15	5	1-360 (step size 1)	2160
10	3, 5, 7, 9		1440
			Total: 3600

~~Section ?? presents results of the~~ In this work, the reference for quantitative analysis of the models' accuracy is the collection of the raw DWM-generated wind fields. Wind sensing accuracy analysis is a straightforward calculation of Root Mean Square Error (RMSE) between the U data points in simulated wind fields and the estimated equivalents. A reference for wake detection accuracy analysis is created via a new classifier trained analogically to the process described in Sec. 2.4, with the only difference being that the training dataset is derived from simulated wind fields, not the estimated ones. Without the bias from the wind field reconstruction, this classifier achieves approx. 99% accuracy under integrated testing. As such, its classifications across all 3600 raw wind fields (with the wind sensing step not performed) are thus taken to be the 'ground truth' reference. Finally, the reference for wake characterisation is obtained by processing a raw simulated wind field: a two-dimensional Gaussian is fitted, moving average filtering is applied, and 'non-impinged' wind slice samples are removed analogically to the process described in Sec. 2.5.

The wind ~~sensing procedure. Section ?? shows the results of the sensitivity study evaluating wake detection performance as relative to the severity of ambient wind conditions. Lastly, in Section ?? the detailed performance of wake detection and wake characterisation together is analysed.~~ sensing accuracy is first analysed by plotting the RMSE across the YZ plane for two example simulations from the testing dataset presenting low and high ambient turbulence. This is shown in Sec. 3.1. The wind-direction-dependent wind sensing and wake detection performance is then analysed via a sensitivity study to U_{amb} and I_{amb}, comparing these models' mean accuracy across the entirety of 10 subsets described above. This is discussed in Sec. 3.2. Sec. 3.3 presents detailed wake characterisation results. This is achieved by visualising framework's performance across selected wind fields from the testing dataset, followed by computing quantitative performance metrics of wake characterisation.

3.1 ~~Wind sensing~~

~~A distribution of RMSE normalised by the corresponding U_{amb} value.~~

~~The accuracy of wind field estimations was first determined by calculating the root mean square error (RMSE) between the estimated and simulated wind speed values. For each simulation,~~

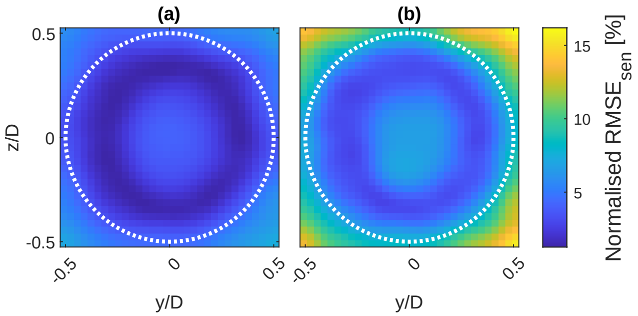


Figure 11. Wind sensing RMSE plotted across YZ slice, normalised by the corresponding U_{amb} value of 10 m/s. (a): I_{amb} = 3%; (b): I_{amb} = 9%. Rotor outline marked with dotted white line.

3.1 Wind sensing: YZ-wise accuracy

Figure 11 compares the ~~RMSE value was calculated separately for each location on the YZ snapshot. Figure 11 shows the~~ YZ-plane distribution of the ~~normalised RMSE for an example simulation . The~~ wind sensing RMSE for low (a) and high turbulence intensity (b). This distribution is obtained by calculating RMSE$_{\mathrm{sen}}$ at every YZ location over the full duration of the simulation n_t:

$$\mathrm{RMSE}_{\mathrm{sen}}(y,z) = \sqrt{\frac{1}{n_t}\sum_{i=0}^{n_t}\left(U_i^{\mathrm{est}}(y,z) - U_i^{\mathrm{DWM}}(y,z)\right)^2} \tag{10}$$

where $U_i^{\mathrm{est}}(y,z)$ and $U_i^{\mathrm{DWM}}(y,z)$ are the YZ-specific U at time i for estimated and DWM wind field, respectively. Both plots show the error to be higher at the rotor centre and in the corners outside the rotor area. The high I_{amb} plot reports significantly higher RMSE, with the critical corner area reaching approximately twice the value of the low-turbulence case. The ring-like pattern of area with lowest error (dark blue in the plot) is repeated across all ~~ambient~~ wind conditions, ~~signifying the wind field reconstruction quality is highest at the radial distance of approx. half of the rotor radius. The flow impacting the middle blade region has the greatest impact on the root bending moments; as a result, the estimation model will be most accurate in this area. The YZ locations outside of the rotor plane yield the highest RMSE due to blade-sensors being unable to meaningfully react to the flow fluctuations from that region.~~ plots for which are not shown here for brevity.

~~The YZ distributions were~~

3.2 Wind sensing and wake detection: ambient conditions sensitivity study

For all 3600 simulations, the YZ distributions of wind sensing RMSE (such as the examples in Fig. 11) are averaged across the entire rotor plane to obtain a single simulation-specific mean ~~RMSE~~ value. Furthermore, to check how the RMSE changes with wind conditions, mean $\overline{\mathrm{RMSE}}_{\mathrm{sen}}$ values averaging all 360 simulations with the same ~~U_{amb} and I_{amb} were calculated. Table 5 shows a comparison of mean RMSE depending on the varying U_{amb} and I_{amb}. It can be noticed how the vast majority~~

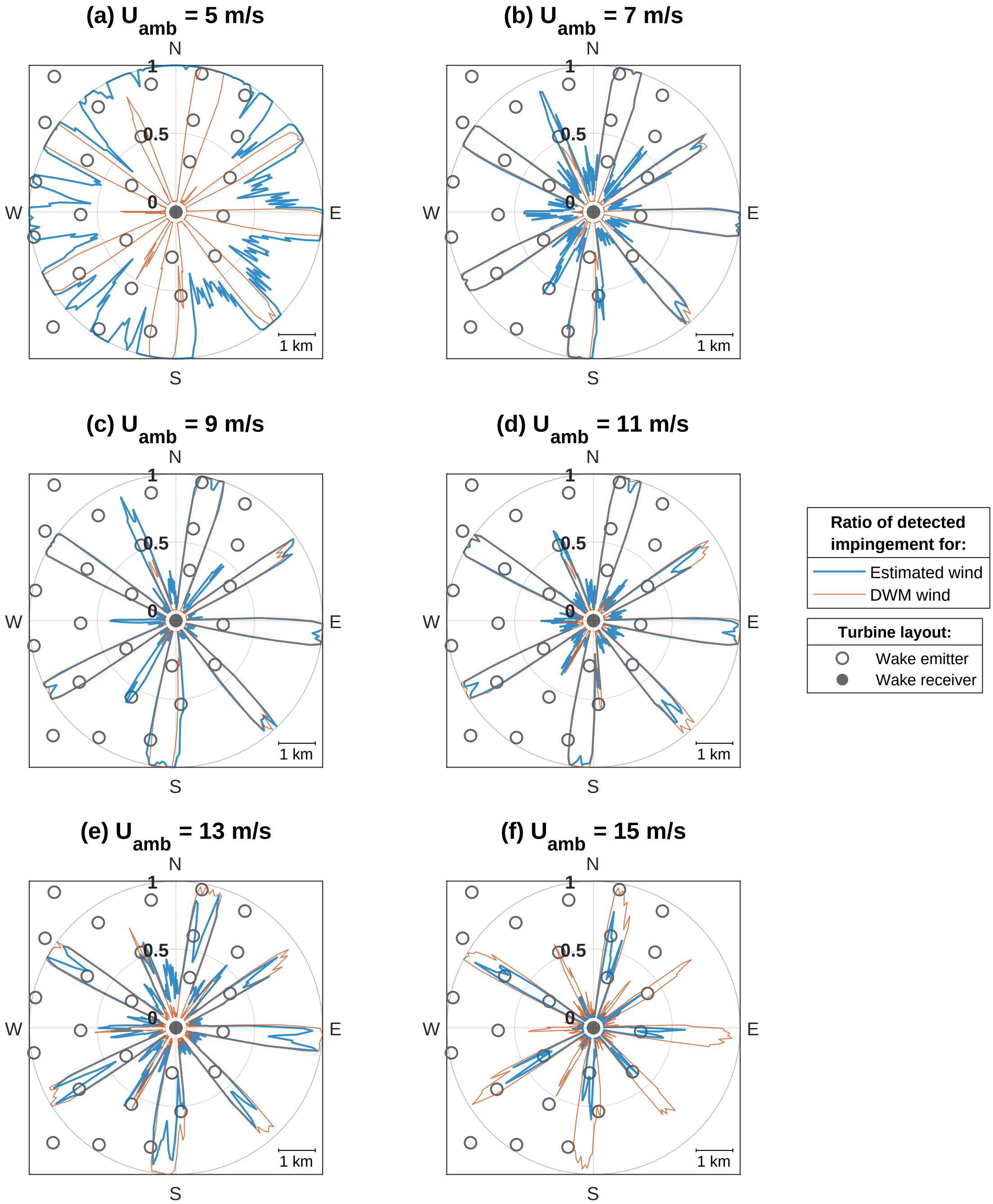


Figure 12. The wake detection model performance tested under a full wind direction range for six U_{amb} values between 5 and 15 m/s. I_{amb} = 5% for all cases. The radial values in blue and orange indicate impingement ratio for a given wind direction. Wind farm layout shown in the background.

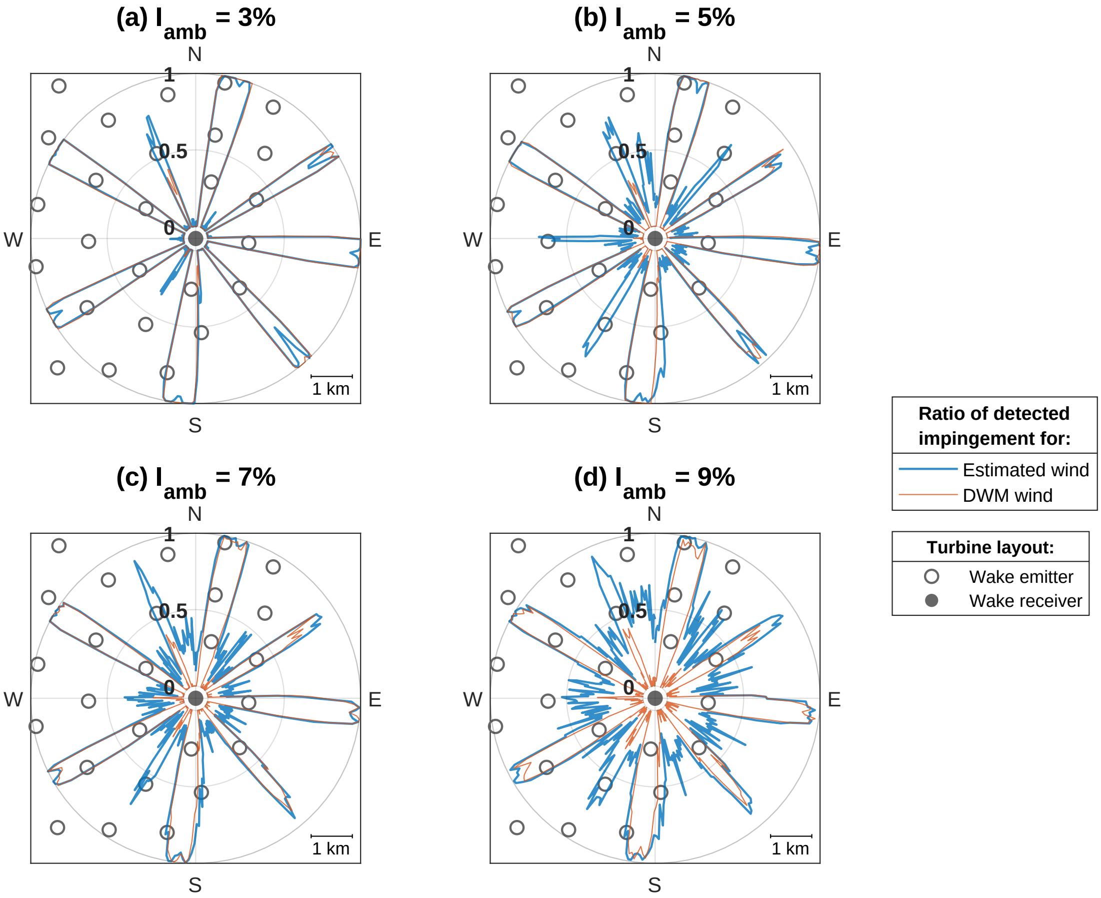


Figure 13. The wake detection model performance tested under a full wind direction range for four I_{amb} values between 3% and 9%. U_{amb} = 10 m/s for all cases. The radial values in blue and orange indicate impingement ratio for a given wind direction. Wind farm layout shown in the background.

~~of cases has a low mean RMSE with the normalised value below 5% of the simulation-specific U_{amb} value, indicating a very good overall wind sensing performance. Increasing the ambient wind speed results in decreasing the normalised RMSE, with the U_{amb} = 5 m/s case showing significantly higher errors than others. On the other hand, the normalised RMSE was found to increase with I_{amb}, with the most turbulent case of I_{amb} = 9% recording the largest normalised mean value of approximately 5.6%.~~ U_{amb} and I_{amb} are calculated. These values are presented in Table 5.

~~Figures ?? and ?? showcase example wind sensing outputs: the former presents two examples of wind fields with well-pronounced wake impingement, while the latter presents two examples of wind fields with no clear impingement.~~

~~A comparison of the simulation wind fields and the estimated equivalents obtained with wind sensing. The wind fields are here presented as a time series of U distribution at the rotor plane at hub height, viewed from above. (a) Partial impingement, U_{amb} = 10 m/s, I_{amb} = 9%; (b) Full impingement, U_{amb} = 10 m/s, I_{amb} = 3%.~~

Table 5. A comparison of ~~averaged RMSE~~ performance metrics in wind ~~speed estimation with~~ sensing and wake detection under varying ~~U_{amb}~~ U_{amb} and ~~I_{amb}~~ I_{amb}. ~~Last column shows values~~ $\overline{\mathrm{RMSE}_{\mathrm{sen}}}$ is normalised by the corresponding ~~U_{amb}~~ U_{amb} value.

~~U_{amb}~~ U_{amb} [m/s]	~~I_{amb}~~ I_{amb} [%]	~~Av. RMSE~~ Normalised $\overline{\mathrm{RMSE}_{\mathrm{sen}}}$ [~~m/sAv. norm. RMSE~~ %]	$\mathrm{RMSE}_{\mathrm{det}}$
5	5	~~0.34~~ 6.7	~~6.74~~ 0.67
7	5	~~0.27~~ 3.9	~~3.86~~ 0.18
9	5	~~0.30~~ 3.3	~~3.30~~ 0.14
11	5	~~0.34~~ 3.1	~~3.09~~ 0.11
13	5	~~0.35~~ 2.7	~~2.69~~ 0.17
15	5	~~0.34~~ 2.3	~~2.28~~ 0.23
10	3	~~0.22~~ 2.2	~~2.22~~ 0.07
10	5	~~0.31~~ 3.1	~~3.12~~ 0.20
10	7	~~0.43~~ 4.3	~~4.26~~ 0.19
10	9	~~0.56~~ 5.6	~~5.57~~ 0.30

~~Looking first at Fig. ??, it can be seen that the major flow fluctuations are well captured in the estimated wind field representations. Some level of smoothing can be seen to occur for~~ For each of the ~~smaller-scale fluctuations in the flow, but this does not pose a major problem in the context of wake detection. In fact, due to the attenuation of less significant turbulent eddies, the wind sensing procedure acts as a spatial filter, allowing to focus only the bigger flow structures like a~~ 3600 simulations, the proportion of samples (YZ wind slices at a point in time) classified as a clear wake impingement (either 'fully impinged', 'partially impinged left', or 'partially impinged right'), is calculated. These simulation-specific impingement ratios are compared to their equivalents in the reference established by the 'ground truth' classifier (see Sec. 3.1). This provides a quantitative measure of the wake detection error for a given wind direction. A single wake detection $\mathrm{RMSE}_{\mathrm{det}}$ value is calculated across the entire wind direction range, for each of the 10 U_{amb} and I_{amb} combinations:

$$\mathrm{RMSE}_{\mathrm{det}} = \sqrt{\frac{1}{n_{\mathrm{wd}}} \sum_{i=0}^{n_{\mathrm{wd}}} (r_i^{\mathrm{est}} - r_i^{\mathrm{DWM}})^2} \tag{11}$$

where: n_{wd} is number of wind directions, r_i^{est} is simulation-specific estimated impingement ratio, r_i^{DWM} is simulation-specific impingement ratio from the raw DWM wind reference.

These values are presented in the rightmost column of Table 5. Figures 12 and 13 show the impingement ratios (both the DWM 'ground truth' and estimated) across the full range of wind directions and under varying U_{amb} and I_{amb} conditions. For each U_{amb} and I_{amb} combination, the plot displays the wind farm layout with the wake 'receiver' turbine in the centre. Polar values visualise the proportion of samples for a given wind direction classified as flow with a clear wake deficit. ~~This aspect can be noticed in comparing the wind fields shown at Fig. ??:~~ These proportions are shown as radial values between 0 (all simulation samples identified as ’no detectable impingement’) and 1 (all simulation samples identified as ’full impingement’, ’partial impingement right’ or ’partial impingement left’).

3.3 Wake characterisation: time-wise accuracy

Figures 14 to 17 provide a detailed performance analysis of the framework across four example wind fields. In each figure, plots (a) ~~portrays high I_{amb}, while~~ and (b) ~~has low I_{amb}. It is apparent that while the some ambient turbulence structure has been smoothed in (a), the wake deficit in both cases is well captured. The wind field shown at Fig. **??** (a) is also noteworthy due to the fact that its ambient conditions of I_{amb} = 9% are outside of the training dataset; however, as seen in the plot, it appears not to be a problem for the developed model, as the estimation accuracy is still high. Overall, these results indicate that this accuracy of wind field estimation is sufficient for wake detection and characterisation.~~ show hub-height horizontal slices of the simulated and estimated U distributions, respectively. The (c) plots display direct wake detection outputs from the CNN’s classification, assigning a value between 0 and 1 to each of the four classes, representing the probability of a sample depicting a corresponding wake impingement case. These probabilities sum up to 1, so a near-perfect class representation has a confidence score near 1, with the other three values near 0. The (d) plots present the outputs of wake characterisation, showing the fitted and filtered lateral wake properties against the DWM reference (see Sec. 3.1).

~~A comparison of the simulation wind fields and the estimated equivalents obtained with wind sensing. The wind fields are here presented as a time series of U distribution at the rotor plane at hub height, viewed from above. (a) No detectable impingement, U_{amb} = 5 m/s, I_{amb} = 5%; (b) No detectable impingement, U_{amb} = 15 m/s, I_{amb} = 5%.~~

~~Figure **??** shows wind sensing results for the wind fields generated~~ Figures 14 and 15 compare the performance under high and low I_{amb}, respectively. They also exhibit different degrees of wake-rotor overlap, allowing to investigate both full and partial impingement scenarios. Figures 16 and 17 show the performance in a scenario with no turbines ~~directly~~ upstream from the ’receiver’ ~~machine. Because of that, no wake impingement should be visible at the time history of U distributions, just the atmospheric fluctuations. The low U_{amb} case shown in (a)reveals a slight discrepancy between the ’Simulation’ and ’Estimated’ subplots. Between 0~~ device. They show the effects of operating under low and high U_{amb} conditions, respectively.

Quantitative metrics for wake characterisation accuracy are computed as shown in Eqs. 12 and 13. Using the reference derived from raw DWM wind fields (see Sec. 3.1), RMSE is calculated over the time series of both lateral wake centre and wake width. Samples classified as ’no detectable impingement’, containing NaN values as discussed in Sec. 2.5, are excluded from the calculation.

$$\text{RMSE}_{\text{char}}^{y_c} = \sqrt{\frac{1}{n_t}\sum_{i=0}^{n_t}\left(y_{c,i}^{\text{filt}}(\text{est}) - y_{c,i}^{\text{filt}}(\text{DWM})\right)^2} \tag{12}$$

$$\text{RMSE}_{\text{char}}^{\sigma_y} = \sqrt{\frac{1}{n_t}\sum_{i=0}^{n_t}\left(\sigma_{y,i}^{\text{filt}}(\text{est}) - \sigma_{y,i}^{\text{filt}}(\text{DWM})\right)^2} \tag{13}$$

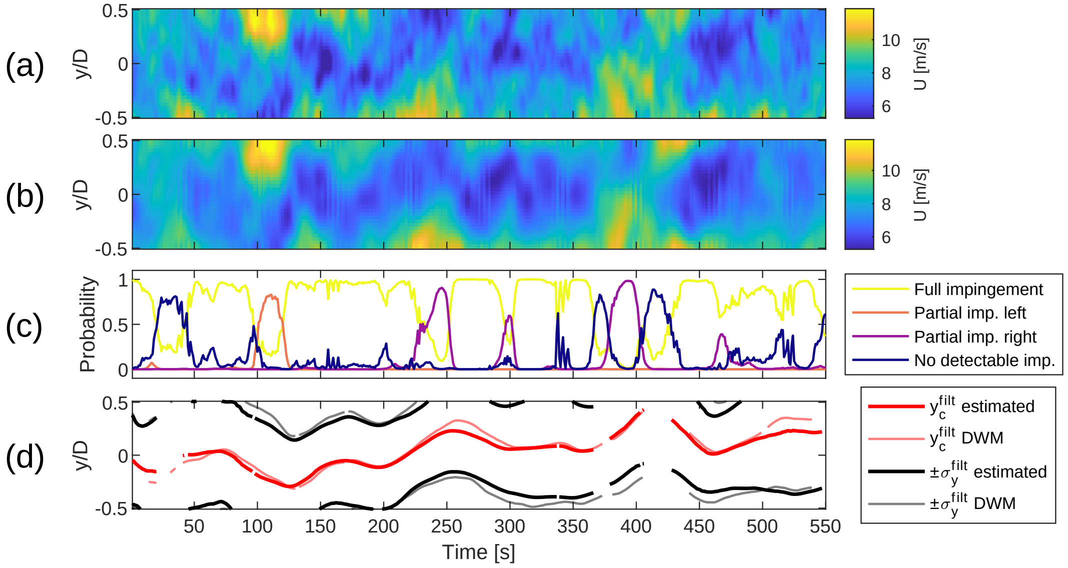


Figure 14. U_{amb} = 10 m/s, I_{amb} = 9%, fully impinged. (a) Horizontal slice of the simulated wind field at the hub height. (b) Horizontal slice of the estimated wind field at the hub height, obtained with the wind sensing model. (c) Classification scores obtained with wake detection model. (d) Fitted wake properties obtained with wake characterisation model next to the reference fitted on the raw DWM wind.

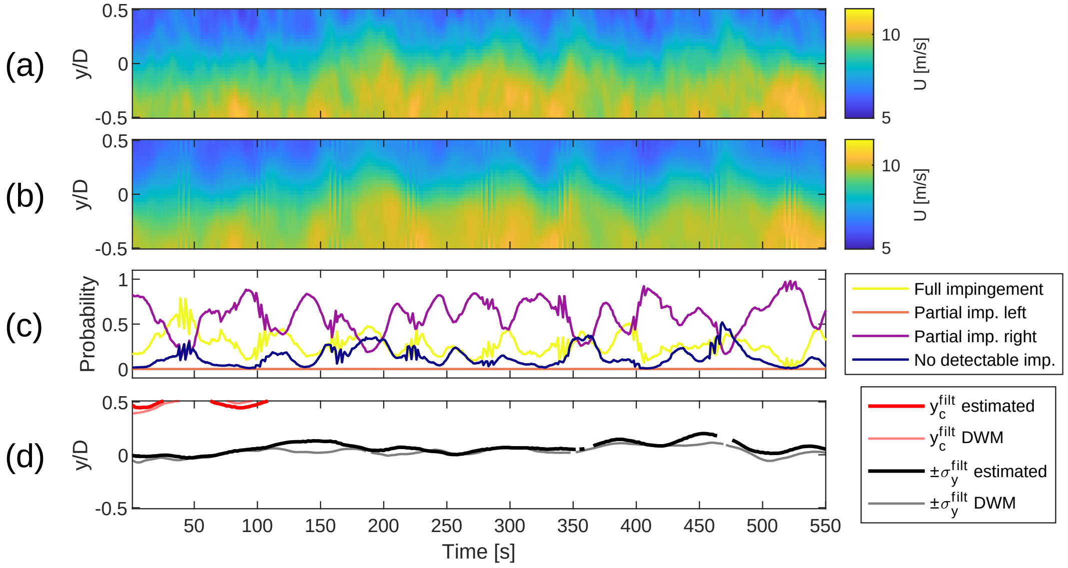


Figure 15. U_{amb} = 10 m/s, I_{amb} = 3%, partially impinged. (a) Horizontal slice of the simulated wind field at the hub height. (b) Horizontal slice of the estimated wind field at the hub height, obtained with the wind sensing model. (c) Classification scores obtained with wake detection model. (d) Fitted wake properties obtained with wake characterisation model next to the reference fitted on the raw DWM wind.

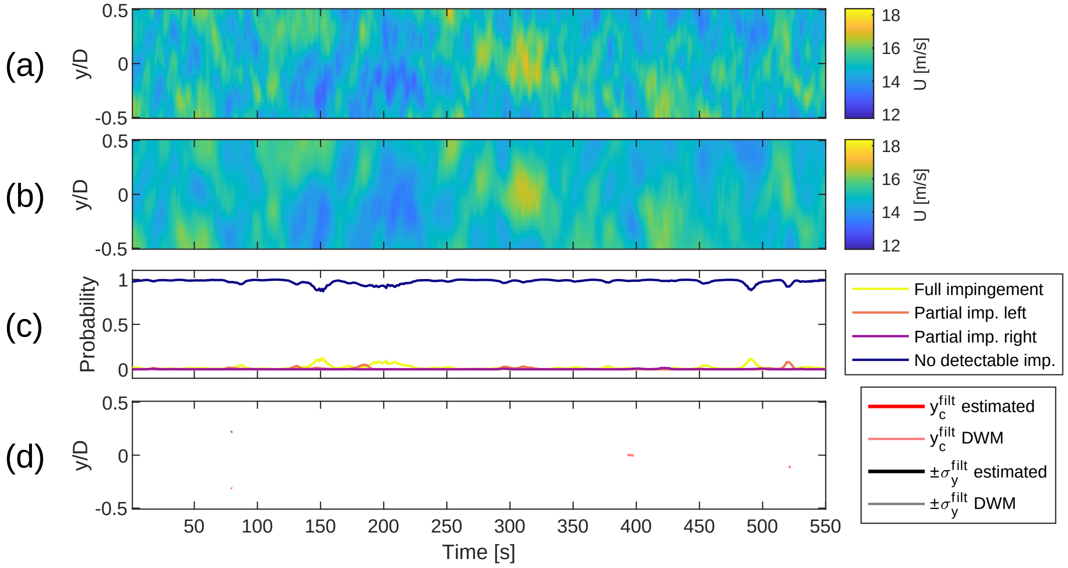


Figure 16. U_{amb} = 5 m/s, I_{amb} = 5%, no impingement. (a) Horizontal slice of the simulated wind field at the hub height. (b) Horizontal slice of the estimated wind field at the hub height, obtained with the wind sensing model. (c) Classification scores obtained with wake detection model. (d) Fitted wake properties obtained with wake characterisation model next to the reference fitted on the raw DWM wind.

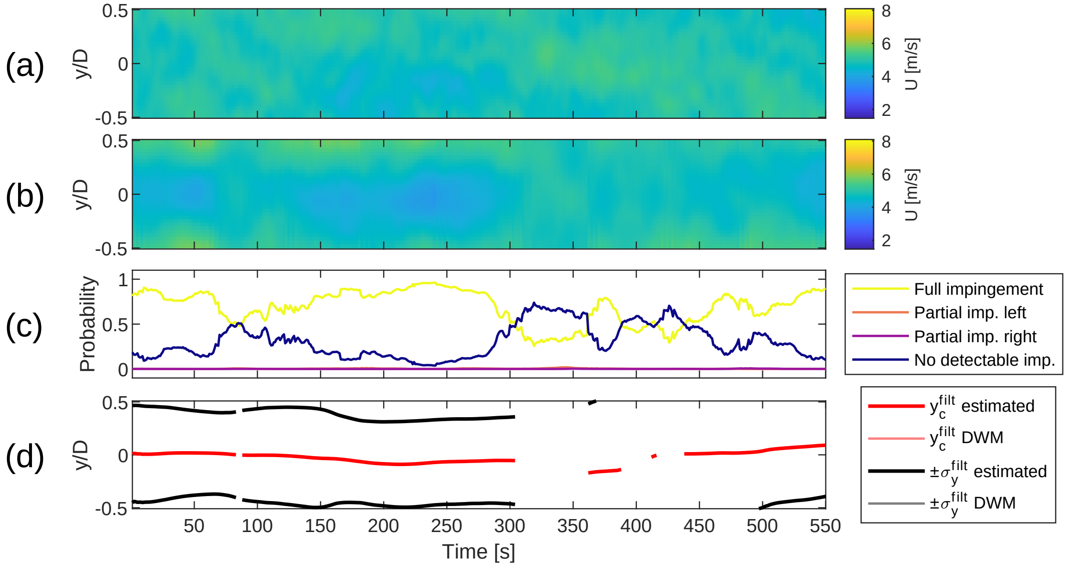


Figure 17. U_{amb} = 15 m/s, I_{amb} = 5%, no impingement. (a) Horizontal slice of the simulated wind field at the hub height. (b) Horizontal slice of the estimated wind field at the hub height, obtained with the wind sensing model. (c) Classification scores obtained with wake detection model. (d) Fitted wake properties obtained with wake characterisation model next to the reference fitted on the raw DWM wind..

Table 6. Performance metrics for wake characterisation.

Impingement case	U_{amb} [m/s]	I_{amb} [%]	$\text{RMSE}^{y_c}_{\text{char}}$ [D]	$\text{RMSE}^{\sigma_y}_{\text{char}}$ [D]
Full	10	9	0.05	0.09
Full	10	3	0.03	0.04
Partial	10	9	0.1	0.1
Partial	10	3	0.03	0.05

where: n_{t} is the simulation duration; $y_{c,i}^{\text{filt}}$(est) and ~~300 seconds, the ’Estimated’plot displays a U deficit in the central region, an effect that is not as prominent in the ’Simulation’ plot. A slight disagreement would not be normally problematic; however, in this case where the flow has little overall turbulence, a subtle deficit like this can’t ’hide’ among other eddies, which could potentially result in classifying the mentioned samples as being wake impinged. The analysis of other low U_{amb} simulations allowed to determine that this is not an isolated example; this phenomenon appears to drive the slightly higher RMSE at low wind speeds recorded in Table 5. Looking at the high U_{amb} case at Fig. **??** (b), it is apparent that as the amount of turbulence increases drastically, effectively dispersing spatially larger fluctuations~~ $\sigma_{y,i}^{\text{filt}}$(est) are estimated y_c^{filt} and σ_y^{filt} at time i, respectively; $y_{c,i}^{\text{filt}}$(DWM) and $\sigma_{y,i}^{\text{filt}}$(DWM) are y_c^{filt} and σ_y^{filt} at time i from the reference fitted on raw DWM wind, respectively.

Table 6 shows the wake characterisation RMSE calculated for example wind fields exhibiting full/partial wake impingement, under high/low I_{amb}. For brevity, only two of these wind fields are visualised in Figs. 14 and 15. Wind fields in which the DWM reference shows primarily ’no detectable impingement’ - such as those in Figs. 16 and 17 - are excluded from the RMSE analysis, as the reference wake characterisation contains mostly NaN values.

4 Discussion

4.1 Evaluation methods

The reference for wake detection is based on the classifications from a detector trained with raw simulated wind fields. This allows to consider the effects of varied wake dispersion under different ambient conditions, and arguably fits this analysis better than a general impingement definition based strictly on inflow angle. Indeed, to the best of author’s knowledge, there isn’t a widely-recognised definition of ‘wake impingement’ (e.g. by means of reduced power output) that could be ~~mistakenly classified as a wake impingement. Good overall performance is therefore seen across the vast range of tested cases, with more significant deviations generally only seen at wind speeds below 7m/s~~used here instead.

Wake characterisation accuracy is evaluated by processing the raw DWM wind with the methods from Sec. 2.5, thus extracting reference wake properties. This approach is preferred over taking the ’ground truth’ from the meandering wake centres applied internally in the DWM model, as turbulent fluctuations in the synthetic wind field — along with additional imposed turbulence - can cause the actual wake experienced by the turbine to deviate from the calculated position. Moreover, this method naturally accommodates the interaction of multiple, combined wakes. A reference derived by fitting a Gaussian profile on a simulated wind is used in other studies (Lejeune et al., 2022).

4.2 ~~Wake detection: sensitivity study~~

~~Figures **??** and **??** show the performance of the CNN-based wake detection framework in a full wind direction range, generated as described in Section **??**. The figures show the sensitivity analysis for varying U_{amb} and I_{amb} in the flow conditions, respectively. For each U_{amb} and I_{amb} value, the plot displays the wind farm layout with the wake 'receiver' turbine positioned in the centre; polar values in the blue visualise the proportion of samples for a given wind direction that were classified as flow with a clear~~ However, it should be noted that using the same methodology to perform estimations and establish reference both, with the only major difference being the source of the wind field, carries a bias. This is especially relevant for calculated wake characterisation metrics; calculated RMSE values could be expected to increase with the use of different reference approach. In future work, we aim to address this by validation with higher fidelity simulations and wind field data.

4.2 Framework performance

Wind sensing

Although there are minor discrepancies for specific wind conditions, the developed wind sensing model presents generally good performance with low error between reference and estimated values. Fig. 11 reveals that the wind field reconstruction quality is highest at the radial distance of approx. half of the rotor radius. This is likely due to the flow impacting the mid-blade region having the largest influence on the root bending moments. The YZ locations outside of the rotor plane yield the highest RMSE due to blade-sensors being unable to react to the flow fluctuations from that region meaningfully. Figures 14 and 15 show that major flow fluctuations are well captured in the estimated wind field representations. Some smoothing occurs for smaller-scale fluctuations, but this does not pose a major problem for wake detection. In fact, the attenuation of less significant turbulent eddies allows the wind sensing procedure to act as a spatial filter, focusing on larger flow structures like a wake deficit. ~~These proportions are presented on the plot as radial values between 0 (all samples in a simulation identified as 'no detectable impingement') and 1 (all samples in a simulation identified as 'full impingement', 'partial impingement right' or ' partial impingementleft').~~ The sensitivity study from Sec. 3.2 shows that the normalised $\overline{\mathrm{RMSE}_{\mathrm{sen}}}$ rises with increasing I_{amb}, which can be attributed to the higher flow complexity in more turbulent wind. The cause for higher error in low U_{amb} cases is investigated in Figure 17. The wind field reconstruction shows a minor U deficit that is not present in the simulated wind field, which is identified by the CNN as a wake impingement. This phenomenon occurs for the majority of wind directions at U_{amb} = 5 m/s (see Fig. 12 (a)), which in turn is the cause for high respective wake detection RMSE captured in Table 5. Looking at the high U_{amb} case at Fig. 17, it is apparent that a more heterogeneous wind field effectively disperses such anomalies. This emergence of 'fake' wakes under low wind speed conditions is likely a result of the training dataset composition, 75% of which shows wake impingement, leaving only 25% for purely ambient turbulence cases. Using a single wind sensing estimator across these cases introduces a bias, as the model is predominantly trained on wake impingement conditions. The misclassification between 'Fully impinged' and 'No detectable impingement' is quantified by the confusion matrix Fig. 9.

~~The wake detection model performance tested under a full wind direction range for six U_{amb} values between 5 and 15 m/s. The radial plot in the blue indicates the proportion of samples classified as wake impingement for a given wind direction. Wind turbine locations indicated with black circles.~~ *Wake detection*

~~The wake detection model performance tested under a full wind direction range for four I_{amb} values between 3% and 9%. The radial plot in the blue indicates the proportion of samples classified as wake impingement for a given wind direction. Wind turbine locations marked with black circles.~~

Examining both Fig. ~~?? and Fig. ??~~12 and 13, most plots ~~depict~~ show the wakes generated by ~~the upstream turbines as~~ upstream turbines to be correctly identified for the corresponding wind directions at the ’receiver’ turbine. There is a distinct difference in the proportion of samples identified as a wake depending on the distance between the ’emitter’ and ’receiver’. For ~~the~~ turbines further away, the proportion of detected impingement is lower due to the wake losing its distinctive shape ~~when developing~~ over a longer distance and interacting with the turbulent ~~ABL~~ atmospheric boundary layer for a longer time. This ~~relation is apparent not only in~~ relationship is evident in both the number of ’impingement’ samples per simulation (~~displayed as the~~ length of the blue~~marker ’tail’) but it is also pronounced in~~/orange peak) and the number of wind directions ~~that the wake was detected for (the~~ for which the wake is detected (width of the blue~~marker ’tail’~~/orange peak).

~~The sensitivity analysis shown in Fig. ?? indicates that the best performance can be observed for U_{amb} values of 7, 9 and 11 m/s. The detected impingement values have a minor noise (up to 0.3 impingement detected for all wind directions) , nevertheless the wakes from all of the nearby turbines are appropriately identified with distinctive values close to 1 occurring for several wind directions each. For U_{amb} =~~ The superposition of wakes does not seem to have any effect on wake detection performance. Wakes from eastern turbines (a single machine upstream) do not give substantially different wake detection results than wakes from deep inside the wind farm to the west. It appears that the primary aspect that decides the wake detection performance is the distance between the ‘receiver’ and ‘emitter’ machines. It is likely a consequence of limiting the training data to one pair of ’emitter’-’receiver’ turbines at a fixed distance.

As shown in Table 5, the wake detection RMSE largely follows the trends from the wind sensing RMSE. This is expected, as one model uses the outputs from the other. The highest error across all cases is reported for ambient wind speed of 5 m/~~s there is a large amount of noise; in fact, for those of the S and E wind directions where there are no turbines upstream, the impingement was detected for approximately half of samples. This phenomenon appears to be caused by the wind sensing deficits observed in very low wind speeds made during the RMSE calculations. The majority of other wind directions report impingement proportions of approximately 1, dropping to lower values for only few angles. For U_{amb} = 13~~ s, which is due to the wind sensing anomaly described above. The wake detection is progressively more accurate with increasing U_{amb} (attributed to dispersion of the anomaly), until the RMSE begins to rise above 11 m/~~s and U_{amb} = 15 m/s, there are small dips visible in the centre of impingement proportion markers, causing them to have a slight ’V’ shape. A manual inspection of the results indicated this is can occur if a large and centered wake covers nearly all of the impinged rotor, presenting similarly to an unimpinged case. Overall, with the exception of U_{amb} = 5 m/s, there is an inversely relationship between the mean ambient wind speed and responsiveness of the wake detection model.~~

~~Figure **??** shows how the wake detection performance is impacted by varying turbulence intensity I_{amb} in the atmospheric flow. The ambient wind speed U_{amb} was here kept at a constant value of 10 m/s. Firstly, it is apparent that increasing the turbulence intensity results in increased noise; this is manifested by having an increasing minimum amount of impingement detected for all of the wind directions for (b), (c) and (d) cases, with values of approx. 0.1, 0.15 and 0.3 respectively. On the other hand, the I_{amb} = 3% case doesn't show significant impingement for turbines located farther away, such as the one directly west of the 'receiver'. This contrasts with the behavior observed in the I_{amb} = 5% case, where more pronounced impingement is reported for turbines at similar distances. For the higher I_{amb} at (c) and (d), the detected wake impingement for turbines at respective wind directions is lower, which could be due to stronger mixing in the atmospheric flow.~~

~~The (d) subplot in Fig. **??** shows the case of ambient turbulence intensity equal to 9%. This is a unique case, as the training data (for both wind sensing and wake detection estimator models) only contained samples with I_{amb} = 3, 5 or 7%. For this reason, the results from this simulation set are a way of presenting how applicable the framework is to working with wind fields that are significantly different from the training data. It can be seen that while the largest impingement proportion is reported at the correct positions, indicating the wakes from nearby turbines correctly, there is now a significant amount of noise for all wind directions.~~

4.3 ~~Wake detection and characterisation: visual results~~

~~Figures **??** and **??** present the performance of the framework in a more detailed manner than discussed in the Sect. **??**. In total, results for six wind fields are showcased, with two subplots each. Firstly, the top plots show direct wake detection outputs from the CNN's classification. For each sample, a value between 0 and 1 is assigned to each of the four classes, representing the probability that a given sample depicts a corresponding wake impingement case. These four probabilities have a total sum of 1, due to which when the sample is a near-perfect representation of a class, the corresponding confidence score is near 1, while the three other values are near 0. Secondly,~~ s. Visual inspection of Fig. 12 (ef) reveals this is a consequence of increased estimation-reference mismatch at the wind directions with high impingement ratio. These wind speeds usually correspond to above-rated operation, where the wakes are less pronounced due to decreased thrust, likely making them more difficult to detect. The wake detection RMSE rises with increasing I_{amb} value; the plots at Fig. 13 show that the main source of the ~~bottom plots report the outputs of wake characterisation, presenting the fitted and filtered lateral wake properties. These were plotted with the simulation wind field shown in the background, providing an easy way to assess the accuracy of the framework.~~

~~Detailed wake detection (top subplots) and characterisation (bottom subplots) results. All three wind fields with clear wake deficit present. (a) U_{amb} = 7 m/s, I_{amb} = 5%; (b) U_{amb} = 10 m/s, I_{amb} = 5%; (c) U_{amb} = 13 m/s, I_{amb} = 5%.~~

~~Cases (a) - (c) in Fig. **??** present varied examples of wind fields where the framework was operating with wakeimpinged flow and it successfully characterised the wake properties. The figure shows how the proportion of successfully detected and characterised snapshots decreases with rising U_{amb}. The time series of fitted wake centre and border is continuous for (a), and progressively becomes disconnected when moving to (b) and (c)~~ error is an increasing 'noise' of some wake impingement detected for all of the wind directions. This is expected, as more turbulent wind fields contain more large-scale ambient eddies that could be identified as a wake. ~~For all three cases it can be seen how the~~

The sample-by-sample classification is presented at (c) plots in Figs. 14 to 17. The CNN's output aligns with the respective wind slice samples; for example, when wake meanders from ~~the~~ centre to the ~~upper~~ Y-positive part of the ~~Y axis~~snapshot, the highest probability transitions from 'Full impingement' to 'Partial ~~impingement~~ imp. left'. ~~Moreover, wake characterisation is not performed for the short periods of~~ This indicates that the instantaneous wake position can be, at least to some extent, tracked even without a dedicated wake characterisation model.

Wake characterisation

Figures 14 and 15 show the sample-by-sample wake deficit characterisation for two example wind fields. Visual comparison between the simulated wind field at (a) plots and characterised wake at (d) plots shows that the wake behaviour is appropriately captured in both cases. The estimated wake properties fit the reference values reasonably well. The time series of fitted wake centre and lateral border are near-continuous for lower I_{amb}, while higher I_{amb} yields a characterisation with few gaps. That is due to the removal of samples classified as 'No detectable impingement' class~~. As seen in (a), due to lowflow turbulence there is almost no wake meandering, resulting in near-constant detection and characterisation outputs. Looking now at (b), it can be seen how the meandering motion is visibly increased, while the framework continues to accurately capture the wake development. The structured deficit dissipates toward the end of the simulation (after t~~ , which happens more often when the deficit shape is distorted by increased ambient turbulence. Table 6 brings further insight, showing quantitative characterisation accuracy metrics for full/partial impingement under low/high I_{amb}. As expected, higher turbulence causes the estimated wake properties to diverge more from the reference, thus increasing the RMSE. For all wind fields, the error is higher for σ_y than for y_c, likely due to the wake centre being more easily identified by the least squares algorithm (it's simply the area of lowest value). Error comparisons between full and partial impingement reveal mostly similar values, the only outlier being $\mathrm{RMSE}^{y_c}_{\mathrm{char}}$, at I_{amb} = ~~450 s); this leads to a brief pause in wake characterisation, followed by a resumption with a significantly broader characterised wake~~ 9%, where the partial wake error is twice that of the full wake. This is a logical outcome, given the increased difficulty of characterisation when wake meandering motion is strong and centre is outside the YZ snapshot.

Figure 16 (d) shows how the 'fake' wake anomaly discussed above results in a characterisation performed on ambient turbulence. This effect is gone when the U_{amb} is higher (see Fig. 17 (d)), which is why the characterisation is missing for the entire wind field except few isolated samples. ~~Moving on to (c) , it can be observed how the ambient turbulence causes the structured deficit to break up, losing the continuity of characterised properties. There are some easy ways to remedy that issue (e.g. Kalman filtering), which provides a basis for future work.~~

~~Detailed wake detection (top subplots) and characterisation (bottom subplots) results. (a) Highly dissipated wake deficit,~~ ~~U_{amb} = 5 m/s, I_{amb} = 9%; (b) Intermittent partial wake deficit, U_{amb} = 15 m/s, I_{amb} = 5%; (c) No wake deficit, U_{amb} = 5 m/s,~~ ~~I_{amb} = 5%.~~

~~Figure **??** presents examples of wind fields which for different reasons proved to be more challenging for the wake detection model. Case (a) at Fig. **??** presents an example of a wind field~~

4.3 Applicability

Overall, the presented performance shows that the developed methodology is a promising solution to the challenging problem of generalised wake impingement estimation. The less accurate wake detection for U_{amb} values of 5 and 15 m/s is less critical after the consideration of the specific application of this work. Firstly, at $U_{\text{amb}} = 5$ m/s the turbine is usually close to its cut-in wind speed (Carrillo et al., 2013), meaning that for many cases, applying flow control would not be necessary. Secondly, the wind farm flow control brings the largest benefits for below-rated operation, where due to ~~high turbulence, the wake deficit is not sufficiently structured and the characterised properties are not continuous. Another consequence of the high turbulence is the more noticeable lag resulting from the moving average filtering; it can be seen that the characterised wake is shifted~~ higher energy extraction and lower wind, the wakes are more pronounced (Scott et al., 2024). These conditions align with the scenarios where the framework demonstrates optimal performance. The wind speeds above 13 m/s (where the framework performance deteriorates) are generally less frequent than lower-wind conditions (Shu and Jesson, 2021), further mitigating the impact of these issues. A similar comment can be made with respect to wake detection RMSE increasing with I_{amb}. Wake steering control brings largest benefits when the wind is less turbulent, which is where the ~~corresponding wind field. Case (b) at Fig. **??** presents an example of a wind field where high atmospheric turbulence causes the wake to intermittently meander into the captured Y range. It can be seen that the CNN classified the majority of samples as 'No detectable impingement' and the wake characterisation was not initiated. At time stamps of approx. 50, 350, 400, 450 and 500 seconds a wake deficit was detected, resulting in a short-term fitting of wake properties. With such a low quantity of identified impingement samples, the characterised wake properties end up being disconnected. Case (c) at Fig. **??** presents a wind field with no turbines upstream from the '~~framework shows best performance.

The presented results also highlight that wake detection accuracy is highly dependent on the training data selection. That indicates that improved performance across a broader range of conditions could be achieved by extension of the training dataset, e.g. by learning the wake deficit shape from various 'receiver' ~~. It is apparent that due to low turbulence and wind speed, the framework mistakenly classified ambient fluctuations as wake impingement. The wind field presented here is the same as the one shown at Fig. **??** (a); a careful analysis allows to determine that the minor deficit noticeable at the 'Estimation' subplot matches the characterised wake centre and border from the Fig. **??**.~~ turbines. One of the testing subsets has I_{amb} equal to 9%, a value outside the training range of 3–7%. While having one of the highest errors in the study, the wakes within the subset are nonetheless detected at correct wind directions. This highlights the framework's ability to produce reasonable results for wind fields with ambient conditions 'unknown' to the estimator. This is highly advantageous for the field application context, where the pre-deployment training procedure would only capture some of the wind effects experienced by the wind farm.

4.4 Current limitations

The present methodology treats each YZ wind snapshot entirely independently, it therefore seems clear that improved performance (and solution to some of the problematic cases which are identified) could be readily achieved by extending the methodology to undertake a post-processing analysis which account for time variations in results. This could, for example, allow for the imputation of gaps that occur in characterised time series for more turbulent wakes (see Fig. 14); moreover, by considering classifications of time-adjacent snapshots, it could help in correct detection of those fully impinged cases which are misidentified as unimpinged. Kalman filtering would seem to be a technique well-suited to these possible methodological extensions.

5 ~~Discussion~~

4.1 ~~Current framework limitations~~

~~Beyond the methodological limitations outlined previously, it~~ It should also be highlighted that the application of a moving average filtering within the methodology results in predictions lagging behind the true wind field. This lag ranges from approximately 12 seconds for ~~U_{amb}~~ U_{amb} = 15 m/s, to approximately 58 seconds for ~~U_{amb}~~ U_{amb} = 5 m/s. These values are a consequence of matching the filtering window width to the wind field's typical meandering time scales (see Eq. (~~??~~7) and (~~??~~8)). In the context of wind farm flow control, these levels of lag aren't necessarily problematic, and the lag may ~~even~~ be removable by ~~undertaking a short-term forecast across the gap. Kalman filteringwould likely be a relatively straighforward way to achieve this. Implementing other approaches such as Long short-term memory~~ introducing methods such as Kalman filtering. Relevant literature (Luo et al., 2024; Zhou et al., 2023) shows that Long Short-Term Memory (LSTM) networks could potentially ~~allow for the forecast to be extended to predict wake locations and meandering behaviour a few minutes ahead~~provide a short-term forecast of the wake dynamics, thus providing an alternative solution. Further research needs to be conducted to investigate these leads.

~~The wake detection results captured at Fig. ?? and Fig. ?? indicate that the higher ambient turbulence and wind speed introduce considerable noise and reduce the size of 'wake markers', meaning that portions of samples were misclassified. Similarly, the wake characterisations obtained during the testing contain several gaps for higher turbulence and wind speed values. However, these issues should be considered less critical after the consideration of the specific application of this work. Firstly, the wind farm flow control brings largest benefits for the below rated operation, where due to higher energy extraction and lower wind, the wakes are more pronounced (Scott et al., 2024). These conditions align with the scenarios where the framework demonstrates optimal performance. Moreover, the wind speeds above 13 m/s (where the framework performance deteriorates) are generally less frequent than lower-wind conditions (Shu and Jesson, 2021), further mitigating the impact of these gaps.~~

Note also that the range of ambient turbulence intensities used for framework testing ~~fell~~ falls between 3% and 9%, and the atmospheric conditions for both training and testing ~~were~~ are considered as neutral. Large offshore wind farms could experience higher levels of turbulence deep within the turbine grid due to multiple wake interactions (Shaw et al., 2022). These simplifications should be acknowledged and other turbulence levels, along with incorporating both stable and unstable ~~ABL~~ boundary layer conditions, analysed in future work. Similarly, turbine yaw offsets ~~were~~ are also not considered in the this work. The current analysis implemented a medium-fidelity wake model, and so further work should also extend this framework to more realistic wake structures, for example those obtained through ~~LES~~Large Eddie Simulations.

4.1 ~~Framework applicability~~

~~Overall, the framework has proved to work well for the majority of tested ambient wind speeds and turbulence intensities. Although there were minor discrepancies for specific wind conditions, the developed wind sensing model presented generally good performance with low error between true and estimated values. The overall average RMSE for wind sensing across all cases was 0.35 m/s, or 5.23% when normalised by respective U_{amb} values. This in turn allowed for good wake impingement analysis based on image recognition. With regards to wake detection, the framework's responsiveness to impingement proved to be dependent on U_{amb} and I_{amb} of the wind field; consequently, the following wake characterisation was sometimes compromised due to wrong classification of some of the samples. The selected bivariate Gaussian wake profile can easily adapt to cases where the wake is 'stretched' due to horizontal sheer or veer. The present methodology and analysis treats each 2D wind snapshot entirely independently, it therefore seems clear that improved performance (and removal of some of the problematic cases which have been identified) could be readily achieved by extending the methodology to undertake a post-processing analysis which account for time variations in results. This could, for example, allow for the imputation of gaps that occur in characterised time series for more turbulent wakes (see Figure **??**); moreover, by considering classifications of time-adjacent snapshots, it could help in correct detection of those fully impinged cases which have been misidentified as unimpinged. Again, Kalman filtering would seem to be a technique well-suited to these possible methodological extensions.~~

~~The presented results also highlight that wake detection performance is highly dependent on the training data, indicating that improved performance across a broader range of conditions can be achieved by enlarging the conditions comprising the training dataset. Overall, the presented performance shows that the developed method for generalised wake impingement estimation is a promising solution to this challenging problem.~~

5 Conclusions

The proposed method for generalised wake impingement detection and characterisation shows good performance for a wide range of tested wind conditions. The framework's effectiveness is sensitive to ambient wind speed and turbulence intensity levels, with ~~optimal~~ best performance observed for wind speeds between 7 and 13 m/s and turbulence intensities ranging from 3% to 7%. Wind field reconstruction reports high accuracy with an all-around RMSE of 3.7% when normalised by mean ambient wind speed. Wake detection correctly responds to a wake presence, and identifying a correct ratio of impingement throughout the simulation duration has an average RMSE of 0.23. Wake characterisation appropriately adapts to the meandering motion, achieving high accuracy; wake centre estimation reports average RMSE of $0.08D$ and $0.03D$ for I_{amb} of 9% and 3%, respectively; wake lateral border estimation reports average RMSE of $0.1D$ and $0.05D$ for I_{amb} of 9% and 3%, respectively. This work provides a baseline concept for generalised wake impingement estimation, which, with further improvements, could greatly contribute to better-informed wind farm flow control. Current limitations are identified, with the most important including: shortcomings of the evaluation methods and anomalies in wind sensing. Next steps for the ~~futher~~ further development of this framework include: validation of the current setup in progressively more realistic environments (including yaw misalignment, higher fidelity wind fields and wakes, or utilisation of field data); an extension of the training dataset (towards accounting more flow phenomena from more turbines at different distances); methodological extension to capture time-variance between individual ~~2D~~ two-dimensional snapshots and opportunities for imputation and forecasting of wake properties. Kalman filtering ~~was~~ is identified as a likely route to extending the framework in this manner.

Appendix A: Convolutional ~~Neural Network~~neural network

The ~~Convolutional Neural Networks (CNNs)~~ CNNs are a specialized type of deep learning method, optimised for working with grid-like data ~~; most commonly 2-D images, but they can also be used for 1-D or multi-dimensional data~~(most commonly two-dimensional images). Their design is perfectly suited for learning spatial features from input data, making them especially effective for image recognition and classification tasks. Due to several key achievements in recent years, they have revolutionised areas such as face recognition, hand writing analysis or autonomous vehicles (Li et al., 2022). In the context of wind energy, some examples of CNN implementation include wind power prediction (Zhu et al., 2017; Harbola and Coors, 2019), early fault detection and classification (Rahimilarki et al., 2022) and others. The input layer of the CNN is the size of the grid-like data it is processing - in a simple example of a grey-scale image, it would be its width multiplied by height. If the task at hand is to classify the image, for example decide on what handwritten digit does the image show, then the output layer would consist of 10 nodes (one for each digit), each one holding the confidence value between 0 and 1. The actual performance happens in the hidden layers located between the input and output layers, and for most CNNs, that is a combination of following:

- *Convolutional Layer*: applies learnable filters (also called kernels) that slide across the input data, only working on a small patch of the data at a time. A dot product of kernel and data at each position creates a measure of how close a patch of data resembles a given spatial feature ~~like an edge or an~~ such as edge or arch. Doing that for the entire input results in a feature map. Performing that operation through multiple iterations with many kernels, progressively optimizing their weights, allows to combine small features into larger ones and ultimately is the key to detecting patterns.

- *Activation Function*: normally following the convolution layers, various activation functions like ReLU (Rectified Linear Unit) introduce non-linearity, enabling the network to learn complex patterns. It is achieved by restricting the connections between the neurons when the weighted sum is too low.

- *Pooling Layers*: reduces the spatial size of feature maps, decreasing computational complexity and helping to prevent overfitting. Max pooling is the most commonly used method of doing so, selecting the maximum value from each patch of the feature map.

- *Fully Connected Layer*: after several cycles of convolution-activation-pooling, this layer connects every neuron in one layer to every neuron in the next, combining features learned by previous layers to produce final predictions.

While designing the neural network, one needs to decide on the number and type of hidden layers implemented in its architecture, as well as on the hyperparameters such as number and size of learnable filters. Moreover, as with all deep learning techniques, the performance of the CNN is heavily dependent on the quality of data used for its training.

Appendix B: ~~A comparison of the univariate and bivariate 2D Gaussian fit~~

~~A particular, univariate case of the 2D Gaussian function occurs when there is no correlation ($\rho = 0$) between Y and Z and when the variations in both axis are assumed to be equal - in that case, $\sigma_y = \sigma_z = \sigma$ and thus the wake is considered to be a circle. The function is simplified as follows:~~

$$f_{G,univariate} = \frac{A}{2\pi\sigma}\exp\left[-\frac{1}{2}\left(\frac{(y_i - y_c)^2 + (z_i - z_c)^2}{\sigma^2}\right)\right]$$

~~A comparison of filtered wake characterisation output time series for univariate and bivariate 2D Gaussian fitting applied. Can be imagined as a view from above - 'true' wind field shown in the background analogically to Fig. **??** and **??**.~~

~~A comparison of wake characterisation performance for a single YZ slice with univariate and bivariate 2D Gaussian fitting applied. Rotor outline marked with black line; fitted circle/oval marked with magenta/cyan solid line; maximum y point (signifying the lateral width parameter) marked with a vertical magenta/cyan line; fitted wake centre for the univariate variant at (y_c, z_c) marked with magenta circle; fitted wake centre for the bivariate variant not visible due to being outside of the snapshot. 'True' wind field YZ slice shown in the background.~~

~~As a result of using a univariate function, the number of parameters to optimize is reduced from six to four: (A, σ, y_c, z_c). Both univariate and bivariate fits were applied to several varied wind fields in search for differences in the wake characterisation performance in this specific application. Generally, it was found out that due to the moving average filtering applied to the fitted wake centre and lateral width time series, the ultimate performance of wake characterisation did not significantly change when switching from one variant of the 2D Gaussian function to another. This effect was captured on an example wind field presented in Fig. **??**. It can be noticed how both estimated wake properties follow approximately the same path. Figure **??** shows a selected YZ slice from the same wind field viewed from the front, where the differences in fitted wake shapes can be easily noticed. The fitted univariate function encapsulates the regular, circular deficit shape, while the bivariate algorithm fits a function that has one of the spreads significantly larger, hence resulting in an ellipse. For that reason, the univariate function is unable to represent cases where the wake is skewed due to turbulence, shear or veer. With that in mind, to allow for a better fit and more accurate wake characterisation results, the bivariate function was implemented in the project.~~

Author contributions. Conceptualisation: Fojcik, Hart, Hedevang. Writing (original draft): Fojcik. Writing (review and editing): Fojcik, Hart, Hedevang. Analysis: Fojcik. Visualisation: Fojcik. Supervision: Hart, Hedevang.

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

Acknowledgements. This research has been funded by Siemens Gamesa Renewable Energy and EPSRC Industrial CASE program (grant no. EP/W522260/1, voucher no. 210039).

References

IEC: International Standard IEC 61400-1: Wind energy generation systems - Part 1: Design requirements, 2019.

Abkar, M. and Porté-Agel, F.: Influence of atmospheric stability on wind-turbine wakes: A large-eddy simulation study, Physics of Fluids, 27, https://doi.org/10.1063/1.4913695, 2015.

Adaramola, M. and Krogstad, P.-A.: Experimental investigation of wake effects on wind turbine performance, Renewable Energy, 36, 2078–2086, https://doi.org/10.1016/j.renene.2011.01.024, 2011.

Ahmed, N., Natarajan, T., and Rao, K.: Discrete Cosine Transform, IEEE Transactions on Computers, C-23, 90–93, https://doi.org/10.1109/T-C.1974.223784, 1974.

Ainslie, J. F.: Calculating the flowfield in the wake of wind turbines, Journal of Wind Engineering and Industrial Aerodynamics, 27, 213–224, https://doi.org/10.1016/0167-6105(88)90037-2, 1988.

Azadkia, M. and Chatterjee, S.: A simple measure of conditional dependence, The Annals of Statistics, 49, https://doi.org/10.1214/21-AOS2073, 2021.

Barthelmie, R., Larsen, G., Pryor, S., Jørgensen, H., Bergström, H., Schlez, W., Rados, K., Lange, B., Vølund, P., Neckelmann, S., Mogensen, S., Schepers, G., Hegberg, T., Folkerts, L., and Magnusson, M.: ENDOW(efficient development of offshore wind farms): modelling wake and boundary layer interactions, Wind Energy, 7, 225–245, https://doi.org/10.1002/we.121, 2004.

Barthelmie, R. J., Pryor, S. C., Frandsen, S. T., Hansen, K. S., Schepers, J. G., Rados, K., Schlez, W., Neubert, A., Jensen, L. E., and Neckelmann, S.: Quantifying the Impact of Wind Turbine Wakes on Power Output at Offshore Wind Farms, Journal of Atmospheric and Oceanic Technology, 27, 1302–1317, https://doi.org/10.1175/2010JTECHA1398.1, 2010.

Bastankhah, M. and Porté-Agel, F.: A new analytical model for wind-turbine wakes, Renewable Energy, 70, 116–123, https://doi.org/10.1016/j.renene.2014.01.002, 2014.

Bertelè, M., Bottasso, C. L., and Schreiber, J.: Wind inflow observation from load harmonics: initial steps towards a field validation, Wind Energy Science, 6, 759–775, https://doi.org/10.5194/wes-6-759-2021, 2021.

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.

Bottasso, C. and Riboldi, C.: Estimation of wind misalignment and vertical shear from blade loads, Renewable Energy, 62, 293–302, https://doi.org/10.1016/j.renene.2013.07.021, 2014.

Bottasso, C. and Schreiber, J.: Online model updating by a wake detector for wind farm control, in: 2018 Annual American Control Conference (ACC), pp. 676–681, IEEE, ISBN 978-1-5386-5428-6, https://doi.org/10.23919/ACC.2018.8431626, 2018.

Cacciola, S., Bertelè, M., and Bottasso, C.: Simultaneous observation of wind shears and misalignments from rotor loads, Journal of Physics: Conference Series, 753, 052 002, https://doi.org/10.1088/1742-6596/753/5/052002, 2016.

Calaf, M., Meneveau, C., and Meyers, J.: Large eddy simulation study of fully developed wind-turbine array boundary layers, Physics of Fluids, 22, https://doi.org/10.1063/1.3291077, 2010.

Carrillo, C., Montaño, A. O., Cidrás, J., and Díaz-Dorado, E.: Review of power curve modelling for wind turbines, Renewable and Sustainable Energy Reviews, 21, 572–581, https://doi.org/10.1016/j.rser.2013.01.012, 2013.

Chamorro, L. P. and Porté-Agel, F.: A Wind-Tunnel Investigation of Wind-Turbine Wakes: Boundary-Layer Turbulence Effects, Boundary-Layer Meteorology, 132, 129–149, https://doi.org/10.1007/s10546-009-9380-8, 2009.

Churchfield, M. J., Lee, S., Michalakes, J., and Moriarty, P. J.: A numerical study of the effects of atmospheric and wake turbulence on wind turbine dynamics, Journal of Turbulence, 13, N14, https://doi.org/10.1080/14685248.2012.668191, 2012.

Cleveland, W. S., Devlin, S. J., and Grosse, E.: Regression by local fitting, Journal of Econometrics, 37, 87–114, https://doi.org/10.1016/0304-4076(88)90077-2, 1988.

Coleman, R. P.: Theory of Self-Excited Mechanical Oscillations of Hinged Rotor Blades, 1943.

Conti, D., Dimitrov, N., Peña, A., and Herges, T.: Wind turbine wake characterization using the SpinnerLidar measurements, Journal of Physics: Conference Series, 1618, 062 040, https://doi.org/10.1088/1742-6596/1618/6/062040, 2020.

Dimitrov, N., Natarajan, A., and Mann, J.: Effects of normal and extreme turbulence spectral parameters on wind turbine loads, Renewable Energy, 101, 1180–1193, https://doi.org/10.1016/j.renene.2016.10.001, 2017.

Dong, L., Lio, W. H., and Meng, F.: Wake position tracking using dynamic wake meandering model and rotor loads, Journal of Renewable and Sustainable Energy, 13, https://doi.org/10.1063/5.0032917, 2021.

España, G., Aubrun, S., Loyer, S., and Devinant, P.: Spatial study of the wake meandering using modelled wind turbines in a wind tunnel, Wind Energy, 14, 923–937, https://doi.org/10.1002/we.515, 2011.

Farrell, W., Herges, T., Maniaci, D., and Brown, K.: Wake state estimation of downwind turbines using recurrent neural networks for inverse dynamics modelling, Journal of Physics: Conference Series, 2265, 032 094, https://doi.org/10.1088/1742-6596/2265/3/032094, 2022.

Frederik, J. A., Doekemeijer, B. M., Mulders, S. P., and van Wingerden, J.: The helix approach: Using dynamic individual pitch control to enhance wake mixing in wind farms, Wind Energy, 23, 1739–1751, https://doi.org/10.1002/we.2513, 2020.

Froude, R. E.: On the Part Played in Propulsion by Differences in Fluid Pressure, Transactions of the Institution of Naval Architects, 30, 390–405, 1889.

Fuertes, F. C., Markfort, C., and Porté-Agel, F.: Wind Turbine Wake Characterization with Nacelle-Mounted Wind Lidars for Analytical Wake Model Validation, Remote Sensing, 10, 668, https://doi.org/10.3390/rs10050668, 2018.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning, MIT Press, 2016.

Hamilton, N., Melius, M., and Cal, R. B.: Wind turbine boundary layer arrays for Cartesian and staggered configurations- Part I , flow field and power measurements, Wind Energy, 18, 277–295, https://doi.org/10.1002/we.1697, 2015.

Harbola, S. and Coors, V.: One dimensional convolutional neural network architectures for wind prediction, Energy Conversion and Management, 195, 70–75, https://doi.org/10.1016/j.enconman.2019.05.007, 2019.

Howland, M. F., Ghate, A. S., Lele, S. K., and Dabiri, J. O.: Optimal closed-loop wake steering – Part 1: Conventionally neutral atmospheric boundary layer conditions, Wind Energy Science, 5, 1315–1338, https://doi.org/10.5194/wes-5-1315-2020, 2020.

Hutchinson, M. and Zhao, F.: Global Wind Report 2023, 2023.

Jensen, N. O.: A note on wind generator interaction, 1983.

Kaimal, J. C., Wyngaard, J. C., Izumi, Y., and Coté, O. R.: Spectral characteristics of surface-layer turbulence, Quarterly Journal of the Royal Meteorological Society, 98, 563–589, https://doi.org/10.1002/qj.49709841707, 1972.

Katić, I., Højstrup, J., and Jensen, N.: A simple model for cluster efficiency, pp. 407–409, 1986.

Keck, R., de Maré, M., Churchfield, M. J., Lee, S., Larsen, G., and Madsen, H. A.: On atmospheric stability in the dynamic wake meandering model, Wind Energy, 17, 1689–1710, https://doi.org/10.1002/we.1662, 2014.

Kim, H., Kim, K., Bottasso, C., Campagnolo, F., and Paek, I.: Wind Turbine Wake Characterization for Improvement of the Ainslie Eddy Viscosity Wake Model, Energies, 11, 2823, https://doi.org/10.3390/en11102823, 2018.

Kim, K.-H., Bertelè, M., and Bottasso, C. L.: Wind inflow observation from load harmonics via neural networks: A simulation and field study, Renewable Energy, 204, 300–312, https://doi.org/10.1016/j.renene.2022.12.051, 2023.

Larsen, G. C., Madsen, H. A., Bingöl, F., and Mann, J.: Dynamic wake meandering modelling, 2007.

Larsen, T. J., Madsen, H. A., Larsen, G. C., and Hansen, K. S.: Validation of the dynamic wake meander model for loads and power production in the Egmond aan Zee wind farm, Wind Energy, 16, 605–624, https://doi.org/10.1002/we.1563, 2013.

Lejeune, M., Moens, M., and Chatelain, P.: A Meandering-Capturing Wake Model Coupled to Rotor-Based Flow-Sensing for Operational Wind Farm Flow Prediction, Frontiers in Energy Research, 10, https://doi.org/10.3389/fenrg.2022.884068, 2022.

Li, Z., Liu, F., Yang, W., Peng, S., and Zhou, J.: A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects, IEEE Transactions on Neural Networks and Learning Systems, 33, 6999–7019, https://doi.org/10.1109/TNNLS.2021.3084827, 2022.

Lio, W. H., Larsen, G. C., and Poulsen, N. K.: Dynamic wake tracking and characteristics estimation using a cost-effective LiDAR, Journal of Physics: Conference Series, 1618, 032 036, https://doi.org/10.1088/1742-6596/1618/3/032036, 2020.

Liu, Y., Pamososuryo, A. K., Ferrari, R., Hovgaard, T. G., and van Wingerden, J.-W.: Blade Effective Wind Speed Estimation: A Subspace Predictive Repetitive Estimator Approach, in: 2021 European Control Conference (ECC), pp. 1205–1210, IEEE, ISBN 978-9-4638-4236-5, https://doi.org/10.23919/ECC54610.2021.9654981, 2021.

Luo, Z., Wang, L., Fu, Y., Xu, J., Yuan, J., and Tan, A. C.: Wind turbine dynamic wake flow estimation (DWFE) from sparse data via reduced-order modeling-based machine learning approach, Renewable Energy, 237, 121 552, https://doi.org/10.1016/j.renene.2024.121552, 2024.

Madsen, H. A., Larsen, G. C., Larsen, T. J., Troldborg, N., and Mikkelsen, R.: Calibration and Validation of the Dynamic Wake Meandering Model for Implementation in an Aeroelastic Code, Journal of Solar Energy Engineering, 132, https://doi.org/10.1115/1.4002555, 2010.

Mann, J.: The Spatial Structure of Neutral Atmospheric Surface-Layer Turbulence, Journal of Fluid Mechanics, 273, 141–168, 1994.

Medici, D. and Alfredsson, P. H.: Measurements on a wind turbine wake: 3D effects and bluff body vortex shedding, Wind Energy, 9, 219–236, https://doi.org/10.1002/we.156, 2006.

Meyers, J., Bottasso, C., Dykes, K., Fleming, P., Gebraad, P., Giebel, G., Göçmen, T., and van Wingerden, J.-W.: Wind farm flow control: prospects and challenges, Wind Energy Science, 7, 2271–2306, https://doi.org/10.5194/wes-7-2271-2022, 2022.

Munters, W. and Meyers, J.: Towards practical dynamic induction control of wind farms: analysis of optimally controlled wind-farm boundary layers and sinusoidal induction control of first-row turbines, Wind Energy Science, 3, 409–425, https://doi.org/10.5194/wes-3-409-2018, 2018.

Nygaard, N. G.: Wakes in very large wind farms and the effect of neighbouring wind farms, Journal of Physics: Conference Series, 524, 012 162, https://doi.org/10.1088/1742-6596/524/1/012162, 2014.

Onnen, D., Larsen, G. C., Lio, W. H., Liew, J. Y., Kühn, M., and Petrović, V.: Dynamic wake tracking based on wind turbine rotor loads and Kalman filtering, Journal of Physics: Conference Series, 2265, 022 024, https://doi.org/10.1088/1742-6596/2265/2/022024, 2022.

Onnen, D., Petrović, V., Neuhaus, L., Langidis, A., and Kühn, M.: Wind tunnel testing of wake tracking methods using a model turbine and tailored inflow patterns resembling a meandering wake, in: 2023 American Control Conference (ACC), pp. 837–842, IEEE, ISBN 979-8-3503-2806-6, https://doi.org/10.23919/ACC55779.2023.10155916, 2023.

Onnen, D., Larsen, G. C., Lio, A. W. H., Hulsman, P., Kühn, M., and Petrović, V.: Field comparison of load-based wind turbine wake tracking with a scanning lidar reference, https://doi.org/10.5194/wes-2024-188, 2025.

Peña, A. and Rathmann, O.: Atmospheric stability-dependent infinite wind-farm models and the wake-decay coefficient, Wind Energy, 17, 1269–1285, https://doi.org/10.1002/we.1632, 2014.

Porté-Agel, F., Bastankhah, M., and Shamsoddin, S.: Wind-Turbine and Wind-Farm Flows: A Review, Boundary-Layer Meteorology, 174, 1–59, https://doi.org/10.1007/s10546-019-00473-0, 2020.

Raach, S., Schlipf, D., and Cheng, P. W.: Lidar-based wake tracking for closed-loop wind farm control, Journal of Physics: Conference Series, 753, 052 009, https://doi.org/10.1088/1742-6596/753/5/052009, 2016.

Rahimilarki, R., Gao, Z., Jin, N., and Zhang, A.: Convolutional neural network fault classification based on time-series analysis for benchmark wind turbine machine, Renewable Energy, 185, 916–931, https://doi.org/10.1016/j.renene.2021.12.056, 2022.

Ronsten, G.: Static pressure measurements on a rotating and a non-rotating 2.375 m wind turbine blade. Comparison with 2D calculations, Journal of Wind Engineering and Industrial Aerodynamics, 39, 105–118, https://doi.org/10.1016/0167-6105(92)90537-K, 1992.

Schreiber, J., Cacciola, S., Campagnolo, F., Petrović, V., Mourembles, D., and Bottasso, C. L.: Wind shear estimation and wake detection by rotor loads — First wind tunnel verification, Journal of Physics: Conference Series, 753, 032 027, https://doi.org/10.1088/1742-6596/753/3/032027, 2016.

Schreiber, J., Bottasso, C. L., and Bertelè, M.: Field testing of a local wind inflow estimator and wake detector, Wind Energy Science, 5, 867–884, https://doi.org/10.5194/wes-5-867-2020, 2020.

Scott, R., Hamilton, N., Cal, R. B., and Moriarty, P.: Wind plant wake losses: Disconnect between turbine actuation and control of plant wakes with engineering wake models, Journal of Renewable and Sustainable Energy, 16, https://doi.org/10.1063/5.0207013, 2024.

SGRE: Siemens Gamesa now able to actively dictate wind flow at offshore wind locations, https://www.siemensgamesa.com/en-int/newsroom/2019/11/191126-siemens-gamesa-wake-adapt-en, accessed: 2024-05-24, 2019.

Shaw, W. J., Berg, L. K., Debnath, M., Deskos, G., Draxl, C., Ghate, V. P., Hasager, C. B., Kotamarthi, R., Mirocha, J. D., Muradyan, P., Pringle, W. J., Turner, D. D., and Wilczak, J. M.: Scientific challenges to characterizing the wind resource in the marine atmospheric boundary layer, Wind Energy Science, 7, 2307–2334, https://doi.org/10.5194/wes-7-2307-2022, 2022.

Shu, Z. R. and Jesson, M.: Estimation of Weibull parameters for wind energy analysis across the UK, Journal of Renewable and Sustainable Energy, 13, https://doi.org/10.1063/5.0038001, 2021.

Simley, E. and Pao, L. Y.: Evaluation of a wind speed estimator for effective hub-height and shear components, Wind Energy, 19, 167–184, https://doi.org/10.1002/we.1817, 2016.

Steinbuch, M., de Boer, W., Bosgra, O., Peters, S., and Ploeg, J.: Optimal control of wind power plants, Journal of Wind Engineering and Industrial Aerodynamics, 27, 237–246, https://doi.org/10.1016/0167-6105(88)90039-6, 1988.

Stevens, R. J. and Meneveau, C.: Flow Structure and Turbulence in Wind Farms, Annual Review of Fluid Mechanics, 49, 311–339, https://doi.org/10.1146/annurev-fluid-010816-060206, 2017.

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

Walker, K., Adams, N., Gribben, B., Gellatly, B., Nygaard, N. G., Henderson, A., Jimémez, M. M., Schmidt, S. R., Ruiz, J. R., Paredes, D., Harrington, G., Connell, N., Peronne, O., Cordoba, M., Housley, P., Cussons, R., Håkansson, M., Knauer, A., and Maguire, E.: An evaluation of the predictive accuracy of wake effects models for offshore wind farms, Wind Energy, 19, 979–996, https://doi.org/10.1002/we.1871, 2016.

Whale, J., Anderson, C., Bareiss, R., and Wagner, S.: An experimental and numerical study of the vortex structure in the wake of a wind turbine, Journal of Wind Engineering and Industrial Aerodynamics, 84, 1–21, https://doi.org/10.1016/S0167-6105(98)00201-3, 2000.

Wu, Y.-T. and Porté-Agel, F.: Atmospheric Turbulence Effects on Wind-Turbine Wakes: An LES Study, Energies, 5, 5340–5362, https://doi.org/10.3390/en5125340, 2012.

Yang, X. and Sotiropoulos, F.: A Review on the Meandering of Wind Turbine Wakes, Energies, 12, 4725, https://doi.org/10.3390/en12244725, 2019.

Zhou, L., Wen, J., Wang, Z., Deng, P., and Zhang, H.: High-fidelity wind turbine wake velocity prediction by surrogate model based on d-POD and LSTM, Energy, 275, 127 525, https://doi.org/10.1016/j.energy.2023.127525, 2023.

Zhu, A., Li, X., Mo, Z., and Wu, R.: Wind power prediction based on a convolutional neural network, in: 2017 International Conference on Circuits, Devices and Systems (ICCDS), pp. 131–135, https://doi.org/10.1109/ICCDS.2017.8120465, 2017.