Predicting the power and loads of wind turbines in waked inflow conditions still presents a major modelling challenge. It requires the accurate modelling of the atmospheric flow conditions, wakes of upstream turbines and the response of the turbine of interest. Rigorous validations of model frameworks against measurements of utility-scale wind turbines in such scenarios remain limited to date. In this study, six models of different fidelity are compared against measurements from the DanAero experiment. The two benchmark cases feature a full-wake and partial-wake scenario, respectively. The simulations are compared against local pressure forces and inflow velocities measured on several blade sections of the downstream turbine, as well as met mast measurements and standard SCADA data. Regardless of the model fidelity, reasonable agreements are found in terms of the wake characteristics and turbine response. For instance, the azimuth variation of the mean aerodynamic forces acting on the blade was captured with a mean relative error of 15–20%. While various model-specific deficiencies could be identified, the study highlights the need for further full-scale measurement campaigns with even more extensive instrumentation. Furthermore, it is concluded that validations should not be limited to integrated and/or time-averaged quantities that conceal characteristic spatial or temporal variations.

© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Wind turbine wakes have been intensively studied ever since the advent of large-scale wind energy exploitation. Still, the modelling of wind turbine wakes and the prediction of power and loads in waked conditions remains one of the greatest challenges in wind energy research [1–3]. Today, a wide range of modelling approaches are being used and actively developed in both industry and academia. Models generally vary in the degree of empiricism, underlying physical assumptions and computational demand. Choosing the appropriate model for a certain application thus entails the assessment of the everlasting compromise between accuracy and cost. While the computational cost is evident to any model, the quantification of model accuracy in the context of wind turbine wakes is arguably difficult. On the one hand, this relates to the complex multi-physics, multi-scale nature of the problem. On the other hand, measurement campaigns capturing the inflow and wake of operating full-scale turbines as well as their structural response and power are costly, technically challenging, and, consequently, rare.

A common approach to evaluate model accuracy and sensitivities are code-to-code comparisons. Examples thereof are several cross-comparisons of actuator line simulations using different large-eddy simulation (LES) frameworks [4–6]. The main motivation for such purely numerical studies is the high comparability of numerical details in a controlled simple setup. Across various numerical schemes and turbulence models, the aforementioned studies showed close agreement of turbine loads and wake statistics. Others compared engineering tools such as dynamic wake meandering (DWM) models to LES results [7–11]. Here, the full availability of flow field data in space and time from the LES allowed for a detailed assessment and tuning of the DWM models.
Such comparisons often rely on the assumption that the high-fidelity LES results are representative of real-world conditions. While this can be the case, Doubrawa et al. [12], for instance, emphasize that LES results remain highly sensitive to the model setup and should therefore be treated with due care.

Wind tunnel experiments offer comparably controlled and reproducible flow conditions while allowing for validations against physical measurements. Over the past two decades, validations have been performed against wake measurements from porous discs [13,14] or model turbines as in the NTNU (Norwegian University of Science and Technology) blind-tests [15,16] or the Mexico [17] and Mexnext [18,19] experiments. Still, typically only a selection of physical features of wind turbine operation can be accurately simulated in wind tunnel experiments, lacking, for example, the important meandering of the wake from turbulent inflow or sufficiently high Reynolds numbers of boundary layer flows.

Comparisons against measurements of full-scale turbines thus remain inevitable for the evaluation of wind farm modelling tools that intend to capture the interaction of the atmospheric boundary layer (ABL), wind turbine wakes and the response of the turbine. Due to its vast availability, the majority of full-scale comparisons have been made to supervisory control and data acquisition (SCADA) data such as power or wind speed measurements from nacelle anemometers [20–24]. Disagreements in the compared quantities are, however, often difficult to interpret because of the lack of information about the inflow or wake. On the other hand, more detailed comparisons including, for example, measurements of the inflow (from met masts or remote-sensing instruments), or wind turbine loads or deformations, are rare.

To date, one of the most detailed field measurement campaigns has been performed at the Scaled Wind Technology Facility (SwIFT) in Lubbock, Texas. The facility features two Vestas V27 turbines, an upstream met mast and a downstream-facing spinning LIDAR mounted on the nacelle of one of the turbines [25,26]. Within Wind Task 31 (WakeBench) organised by the International Energy Agency (IEA), a comprehensive benchmark study has been performed, comparing single wake simulations against measurements from SwIFT [12]. The participating models include analytical wake models, Reynolds-averaged Navier-Stokes (RANS) approaches, DWM models, as well as LES. In neutral conditions, the model performance was found to be closely correlated with the model fidelity. Conversely, in stratified conditions, the RANS models outperformed the LES models because an accurate prescription of the complex inflow conditions was more straightforward. The authors generally concluded that the main challenge for wake models and their experimental validation relates to the simulation of the inflow. Furthermore, the largest discrepancies between the time-resolved models originated from the differences in the wake meandering. Both aspects demand for future measurement campaigns that include highly resolved comparisons of the flow both up- and downstream of the turbine, as well as more detailed load measurements.

Over the past decade, several wind turbine modelling benchmarks have been conducted within IEA Wind Task 29 (hereafter referred to as Task 29). Previous phases of the task were concerned with the aforementioned Mexico and Mexnext experiments. The final stage (Phase IV) focused on comparisons to full-scale measurements from the DanAero experiment [27]. The DanAero experiment was conducted at the Tjæreborg wind farm in Denmark between 2007 and 2009. The experiment featured a met mast within the farm as well as a large range of sensors on one of the eight 2.3-MW wind turbines. Due to the large range of
simultaneously measured quantities, the DanAero measurements still present a unique experimental database that serve various validations studies, the most recent ones being part of Task 29 [28–34]. While focusing on different aspects, most previous studies using the DanAero database considered undisturbed inflow conditions. This relates to both wind turbine loads and wake properties. The dataset thus remains largely unexploited when it comes to waked inflow conditions. Indeed, to the authors’ knowledge, full-scale validations of wind turbine loads in waked inflow are generally untouched. The presented paper aims to close this gap by providing a first comparison of two wake inflow benchmark cases organised within Task 29. More specifically, we compare six different numerical models of three different fidelities (DWM, RANS and LES) against DanAero measurements in a full wake and partial wake case, respectively. In both cases, the turbine of interest is subject to the wake of a single upstream turbine.

The rest of the paper is organised as follows. In Sec. 2 we provide an overview of the Tjæreborg wind farm, the setup of the DanAero experiment and the specific measurements used for comparison in this study. Furthermore, the particular flow cases of the benchmark and the definition of the benchmark itself are described. Sec. 3 outlines the fundamentals of the participating models and the respective numerical setup. Results are presented in Sec. 4 and 5 further discussed in Sec. 6. Final conclusions and an outlook of future work is given in Sec. 7.

2. Measurements and benchmark definition

The Tjæreborg wind farm is located on the west coast of Denmark about 10 km southeast of the town of Esbjerg. The eight 2.3-MW NM80 turbines are placed in two rows of four turbines each, parallel to the south-southwest-facing coastline. The row-normal distance to the shore measures about 600 m referring to 7.5 rotor diameters $D$, with $D = 80$ m. The turbine hub height is $H = 57$ m. The surrounding topography is mostly flat agricultural land with low roughness. A schematic of the farm is given in Fig. 1. In the following we provide a brief overview of the experimental data used in this study. For further details, we refer to the original project reports [27,35].

2.1. Atmospheric inflow

A characterisation of the inflow is provided by measurements from the met mast shown in Fig. 1. The location of the mast relative to the turbine position and inflow direction ensures mostly undisturbed inflow measurements. Still, induction effects of the surrounding turbines or influences of the wake of WT1 can obviously not be ruled out. For this study, we make use of the velocity measurements from the sonic anemometers mounted at three heights $z = \{17, 57, 93\}$ m. The sonics (Metek Sonic USA1) provide three velocity components sampled at 35 Hz. Additional temperature measurements along the mast were only used in the initial selection of suitable cases. The selected cases (further discussed in Sec. 2.5) refer to near-neutral atmospheric stability. Temperature is therefore neglected in all participating models and thus not required to prescribe the inflow conditions in the simulations.

2.2. Blade forces

The forces acting on four airfoil sections of one of the blades of WT2 are obtained from local pressure measurements. As part of the DanAero project, 64 pressure taps were embedded along chord-wise rows at four radial positions $r = \{13, 19, 30, 37\}$ m on both the pressure and suction sides of the blade. The integration of the pressure along the rows provides an estimate of the local force acting on the respective airfoil section. Further details upon the setup, calibration and postprocessing can be found in Madsen et al. [27].

2.3. Blade inflow

Pitot tubes were mounted at four radial positions of the blade to measure the local inflow velocities. The Pitot tubes provide three velocity components measured at 35 Hz at about one chord length upstream of the leading edge. Unfortunately, only one of the Pitot tubes (at $r = 20.3$ m) was functional during the two bins considered in this study. For further details on the experimental set-up, we refer to Ref. [27].

2.4. SCADA data

In addition to the instrumentation of the DanAero experiment, the ordinary SCADA data of WT2 were recorded at 35 Hz. For WT1, only 10-min averages are available for the considered time period. For the presented benchmark comparison, we compute the mechanical power, $P$, from the electrical power logged by the SCADA system. The mechanical and electrical efficiency required for the conversion are obtained from torque measurements on the main shaft provided in Ref. [36].

2.5. Case selection

Under consideration of multiple criteria, two 10-min bins, hereafter referred to as Case A (measured on July 21, 2009, 11:20h–11:30h) and Case B (July 21, 2009, 13:10h–13:20h), are selected from the available data to serve as individual benchmark cases. The target atmospheric criteria are near-neutral stability, suitable mean wind direction and small long term transients in the wind direction and speed, where long term refers to time scales in the order of minutes. The latter is primarily motivated by the fact that the turbines in the simulations are set to operate at a constant rotational speed and yaw corresponding to the respective means measured during each bin. Large-scale deviations from the mean velocity or wind direction of each bin would thus imply overly long periods of sub-optimal operational conditions in the simulations when compared to the actual turbine. A complementary summary of the

![Fig. 1. Schematic of the Tjæreborg wind farm (background image © Google Maps 2021). The wind turbine locations are marked by red dots. For the cases discussed in this work the main wind direction is southwest-west indicated by $\omega_0$. WT1 refers to the upstream turbine inducing the wake. WT2 (blue dot) refers to the equipped turbine. The met mast (MM, orange square) is located between the two turbines with a small lateral offset with respect to the wake of WT1. The connecting dashed line between WT1 and WT2 refers to a wind direction of $\theta = 257^\circ$.](image)
characteristic case parameters discussed in the following is provided in Table 1. Mean quantities are denoted by overbars ($\bar{\cdot}$).

The atmospheric stability is evaluated based on the stability parameter

$$\frac{z}{L} = \frac{\phi_h}{\phi_m}$$

(1)

where $L$ is the Obukhov length, $\text{Ri}$ the gradient Richardson number (computed at $z = 37$ m using the velocity and temperature measurements of the sonics) and $\phi_h, \phi_m$ the Monin-Obukhov functions approximated by the formulation by Wilson [37]. With $z/L$ of $-0.08$ and $-0.12$, respectively, the two cases are found to be near-neutral with a tendency to weakly unstable conditions.

As for the mean wind direction, at hub height the selected Case A features $\bar{\theta} = 251.08^\circ$, resulting in a partial wake impact on WT2. In Case B, the mean wind direction is almost exactly aligned with the two turbines ($\bar{\theta} = 257.02^\circ$) serving as a full-wake scenario. The time series of the magnitude of the horizontal wind velocity, $u_h$, and wind direction measured by the sonics are shown in Fig. 2.

In Case A, the velocity in the upper two measurement locations exhibits relatively small fluctuations. The mean turbulence intensity $Tl = \sqrt{2/3 \text{TKE}}/u_h$ (TKE being the turbulence kinetic energy) at hub height amounts to 6.98%. The wind direction plots reveal a notable veer across the rotor swept area with maximum values $\Delta\theta_h = \theta_{93} - \theta_{17}$ of up to 38°. The mean veer amounts to 10.53°. For Case B, similarly low large-scale fluctuations are found while the hub height turbulence intensity is higher with $Tl = 12.10\%$. The mean veer measures $11.14\%$. A similar mean shear exponent, $\alpha$, of 0.093 and 0.098, respectively, is found for the two cases. Another crucial difference between the two cases is the applied control strategy. In Case A the rotational speed $\omega$ of both turbines is dynamically controlled by the default generator torque controller.

For WT2, for instance, this results in a standard deviation of $\omega$ of $\sigma(\omega)/\omega = 0.12$. On the other hand, in Case B, as part of the DanAero experiment the controller of WT2 was set to operate at a constant rotational speed to facilitate comparisons like the one presented here.

2.6. Benchmark definition

The main objective of the benchmark, as defined within Task 29, is the model comparison of wind turbine power and loads under waked inflow conditions. The choices made in the definition of the benchmark thus state a compromise between the comparability among the models as well as the comparability to the experiment.

Previous studies have shown that the accurate reproduction of the measured inflow conditions is among the most crucial aspects when benchmarking against full-scale data [12]. At the same time, the optimal method to impose the inflow can vary between different model types. In this comparison, a common turbulent inflow field was generated for both cases and provided to all participants. The inflow field was generated using the open-source toolkit PYCONTURB [38]. Turbulence boxes generated by PYCONTURB are based on the Kaimal spectrum and enforce exponential coherence in the stream-wise direction while following multiple user-defined constraints; in this case, the time series of the velocities measured by the three sonic anemometers. The main motivation to provide a common turbulence box was to reduce differences between the simulation results arising from different choices in the inflow generation. Furthermore, given the short length of only 10 min per case, it appeared desirable to explicitly replicate the measured inflow as close as possible, as opposed to imposing turbulence that only matches certain statistics. The only choice left to the participating modellers was to uniformly scale the imposed turbulence box. This option was motivated by the fact that synthetic turbulence imposed at the inlet of LES evolves/decays while propagating downstream. The scaling factor can thus be adjusted such that the turbulence statistics at the reference points (the locations of the sonics) more closely match the statistics of the imposed inflow [8,39,40]. The inflow planes measure 7.5 $D$ (800 m) in the lateral and vertical directions, respectively, with a spatial resolution of $\Delta y = \Delta z = 0.0625$ $D$ (5 m). The temporal resolution is $\Delta t = 0.1$ s. Note that for a complementary comparison of the impact of the synthetic inflow in LES, one model was run with the provided inflow and with a precursor-generated inflow, respectively. Furthermore, the RANS model inevitably applied the time-averaged velocity and TKE of the inflow planes as inflow conditions. Additions to the commonly shared turbulence boxes were also

| $\bar{\theta}$ | $\Delta\theta_h$ | $\frac{z}{L}$ | $\bar{u}_h$ | $\bar{Tl}$ | $\bar{\omega}$ |
|---------------|----------------|-------------|-------------|------------|--------------|
| $\bar{\theta}$ | $\Delta\theta_h$ | $\frac{z}{L}$ | $\bar{u}_h$ | $\bar{Tl}$ | $\bar{\omega}$ |

Note that the $\bar{\omega}$ values represent the rotational speed in the horizontal and vertical directions, respectively, with a spatial resolution of $\Delta y = \Delta z = 0.0625$ $D$ (5 m). The temporal resolution is $\Delta t = 0.1$ s. Note that for a complementary comparison of the impact of the synthetic inflow in LES, one model was run with the provided inflow and with a precursor-generated inflow, respectively. Furthermore, the RANS model inevitably applied the time-averaged velocity and TKE of the inflow planes as inflow conditions. Additions to the commonly shared turbulence boxes were also.
employed in the DWM models to capture large-scale wake meandering; see Sec. 3 for more information.

Other choices regarding the numerical set-up (grid resolution, domain size, boundary conditions, turbulence models, etc.) were left to modellers. A schematic of the case set-up including the coordinate definition is given in Fig. 3. Both cases will be discussed within this coordinate system (as opposed to aligning \( x \) with the respective mean wind direction of each case), with \( u \) being the stream-wise, \( v \) being the lateral and \( w \) being the vertical velocity component.

As for the turbine modelling, the participants were provided with the standard aerodynamic data required for common actuator-type models (airfoil polars, chord and thickness distributions, etc.). Moreover, as mentioned earlier, the turbines were to be modelled with a fixed \( \alpha \) and \( \gamma \) angle referring to the mean of the respective measured values of the cases; see Table 1. This way, additional differences between the models due to different choices in the controller settings, or the controller response to differences in the inflow could be avoided. For the same reason, aeroelastic effects were disregarded in all turbine models. However, note that fully resolved aeroelastic simulations of the NM80 turbine showed rather small effects of the blade deflections on the thrust, power and wake of the turbine due to the high stiffness of the rotor [30,31].

3. Numerical models

Six different models of three different fidelities participated in the benchmark; see Table 2.

In the following sections we provide a brief description of the models and the respective setup of the simulations discussed in this study. Further numerical details can be found in A.

3.1. RANS-OpenFOAM

OpenFOAM is an open-source finite volume solver, widely used within academia and industry. In this case, the OpenFOAM version 4.1 is employed for stationary RANS simulations. The utilized setup has previously been compared to measurements and LES results of wind turbine wake interaction problems [41–44].

The SimpleFoam steady-state solver, based on the SIMPLE (Semi-Implicit Method for Pressure Linked Equations) algorithm, is used. The convective terms are discretized with a bounded, second-order accurate, linear upwind scheme. The Realizable k-\( \epsilon \) model [45] with standard coefficients is chosen as the turbulence model. At the inlet, a neutral ABL profile is defined by imposing the velocity at hub height and the terrain roughness length \( z_0 \), such that the target ambient turbulence intensity is achieved [46]. According to that adjustment, \( z_0 = 0.02 m \) is prescribed in the wall function applied at the bottom boundary condition. The lateral faces are periodic and at the outlet a Neumann condition is imposed. On the top face the velocity is fixed with the value obtained in the inlet profile at the top height. The turbine is modelled using an actuator disk model (ADM) that incorporates airfoil data [46,47]. The forces are calculated at nodes distributed on multiple lines that cover the entire rotor swept area [48,49] and applied to the cells using an isotropic Gaussian convolution kernel. The calculated forces are corrected with the tip correction proposed by Shen [50,51].

3.2. DWM-FAST.Farm

FAST.Farm [52] is an open-source solver for wind farm simulations developed by the National Renewable Energy Laboratory (NREL). The tool consists of three main modules: 1) Each wind turbine is modelled using an instance of OpenFAST, allowing for full aero-hydro-servo-elastic simulations based on blade-element-momentum (BEM) aerodynamics. 2) The wakes of individual turbines are solved by a wake dynamics module, which tracks different wake deficit planes, convects them, and accounts for their dynamics (e.g. diffusion, expansion) by solving the thin shear layer approximation of the Navier-Stokes equation in polar coordinates in the meandering frame of reference with closure accounted for via an eddy viscosity model. 3) The ambient and wind array effect module is responsible for handling the turbulent background flow of the wind farm, merging the wake deficits of individual turbines and providing the velocities requested by the two other modules. FAST.Farm is based on the original DWM model from the Technical University of Denmark (DTU) [53], but with advanced corrections. The version of FAST.Farm employed here does not have a wake-added turbulence model implemented, which was still under development at the time of writing.

The FAST.Farm simulations used the original grid size from the turbulence boxes in the vertical and lateral directions (5 m in both directions). The longitudinal grid size was set to 4.9 m and 4.2 m for cases A and B, respectively. The time step of OpenFAST was set to 0.01 s and the ones of FAST.Farm were set to 1 s for the low-resolution background flow and 0.1 s for the high-resolution flow near the turbines. The polar grid of WakeDynamics used 50 radial nodes spaced by 4 m. Default values for the wake evolution and
meandering parameters that are adjustable within FAST.Farm were used.

The lack of spatial coherence in the two transverse (lateral horizontal and vertical) directions in the commonly shared turbulence boxes results in too little wake meandering in DWM models. To get around this, NREL used modified versions of the commonly shared turbulence boxes, also generated by Pycnturb, but which include exponential coherence in the two transverse directions while preserving the same stream-wise velocities as the commonly shared turbulent boxes. In Case B, the mean and standard deviation of the longitudinal velocity from the turbulence box were also scaled by 3% and 15%, respectively, to better match the results from other participants.

3.3. DWM-DTU

The DWM-DTU model stands for the DWM model implemented in the aeroelastic tool HAWC2 [54]. The model underwent various improvements since its first presentation in 2003 [55] including the development of a wake-added turbulence model [56,57] based upon preceding full-scale validation studies, see, for example, [58]. The current version computes the initial wake deficits based on the BEM wake equations and simulates the wake development downstream with an eddy viscosity model. The large-scale meandering of the wake is driven by an additional turbulence box advecting the wake deficit in the lateral and vertical directions. Details can be found in Ref. [53] along with further calibration and validation studies presented in Refs. [27,59].

In this study, a time step of 0.1 s was used for both HAWC2 and the integrated DWM-DTU model while a refined time step of 0.02 s was used in the simulations of blade forces and power. For the large-scale meandering, a Mann box with a turbulence intensity equal to the ambient turbulence was used with a lateral and vertical extent of 8 D and a grid size of 1 D.

3.4. LES-elbe

The efficient lattice Boltzmann environment (elbe) was originally developed at the Technical University of Hamburg [60] and later extended for wind energy applications at Uppsala University [6,61]. The underlying LES lattice-Boltzmann solver is efficiently parallelised on graphics processing units (GPUs) using Nvidia’s CUDA toolkit.

In the presented simulations, we employ a D3Q27 lattice and the parametrised cumulant collision model [62] along with a standard Smagorinsky turbulence model with a model constant of $C_S = 0.1$. The velocity at the inlet is prescribed with an equilibrium wet-node approach [63]. A symmetry boundary condition is applied at the top of the domain (simple bounce-forward; see, e.g. Ref. [64]). At the bottom, we set the wall shear stress using a simple bounce-back scheme coupled to the wall modelling approach presented in Ref. [65]. The domain is periodic in the lateral direction. A linear extrapolation boundary condition is applied at the outlet as well as a viscous sponge layer that extends 1 D upstream into the domain. The grid in the turbine vicinity and wake region is refined with a nested grid refinement zone. For the interpolation of distribution functions between the fine and coarse grid we employ a compact second-order interpolation [66,67]. The turbines are modelled using an ALM. The implementation closely follows the original description by Sørensen and Shen [68]. Further LBM-specific details of the model can be found in Refs. [6,61,69]. The simulations are run at a Mach number of $Ma = 0.1$, referring to a time step of $\Delta t = 0.018$ s on the finest grid.

3.5. LES-ElipSys3D

ElipSys3D [70–72] is a general purpose flow solver developed at the Technical University of Denmark. It solves the incompressible Navier-Stokes equations in structured curvilinear coordinates following the finite-volume method on collocated grids. The code is parallelised using the message passing interface (MPI) library and applies a multi-grid and grid sequencing approach to accelerate the computations. Multiple numerical schemes are available, however, here the coupled momentum and pressure-correction equations are solved using an improved version of the SIMPLE algorithm [73], and the convective terms are discretized with a fourth-order accurate central difference scheme. The solution is advanced in time by an iterative, implicit second-order scheme and six sub-iterations are used. A modified Rhie-Chow interpolation avoids velocity-pressure decoupling [74]. Sub-grid scales (SGS) are modelled using Dearoff’s turbulence model [75,76]. The inflow and outflow faces obey Dirichlet and Neumann conditions, respectively, whereas the top and bottom faces are of the symmetry type. The lateral faces are periodic to allow for cross-flow. The predefined inflow profile is prescribed at the inflow plane, which is 3D upstream of WT1 in the region where the prescribed turbulent inflow is superimposed. The corresponding power law is enforced. The ALM with viscous core correction by Meurer Forsting et al. [77–79] represents the turbine blade forces, where forces are distributed in the domain by an isotropic Gaussian kernel with a smearing length scale of 2.5A to minimise the angle-of-attack error [80]. A total of 15 min are simulated, with statistics taken over the last 10 min.

3.6. LES-OpenFOAM

The OpenFOAM framework (version 2.3.1), in conjunction with the Simulator for Wind Farm Applications (SOWFA) project [81], developed by NREL, are used. SOWFA was created for LES computations of wind turbines in ABLs and has been widely used and validated [51,81–83]. The atmospheric solver used in SOWFA is called ABLSolver and it is developed out of the buoyantBoussinesqPimpleFoam code from OpenFOAM. It is a transient solver for turbulent flow of incompressible fluids and considers Boussinesq approximation for buoyancy effects. The convective terms are discretized with a second-order accurate linear scheme. For this study, only neutral conditions are considered and a capping inversion starting from 700 m above ground is fixed. The inclusion of the capping inversion follows the recommendation by Ref. [81]. A one-equation eddy-viscosity (SGS) model is chosen [84]. The turbines are modeled using the ALM [68].

Two different domain configurations are used for the two inflow cases LESUU and LESULPC, respectively. A detailed description can be found in Tab. 4.

In LESULPC the inflow is generated by a precursor simulation of 5 h (18000 s), following the recommendations in Ref. [82] to obtain a quasi-steady state. In the next hour and 10 min, the values at the corresponding inlet faces and the pressure gradient forcing are recorded for each time step, in order to be used in the farm simulation. The precursor is generated with periodic boundary conditions on the inlet, outlet and lateral faces. The flow is pressure-driven and this forcing is adjusted every time step to get the desired average velocity and direction on a plane at hub height. The top face is slip and on the bottom face a wall function is applied. If the corresponding power law is enforced. The inclusion of the capping inversion follows the recommendation by Ref. [81]. A one-equation eddy-viscosity (SGS) model is chosen [84]. The turbines are modeled using the ALM [68].

In LESULPC the inflow values generated with the turbine box. The flow is pressure-driven and the forcing values are copied from LESUU.
4. Case A - partial wake

In the following, we present the results from Case A with WT2 being subjected to a partial wake inflow. The discussion follows a sequential downstream order starting from the inflow, on to the power and thrust of the upstream turbine and finally the wake inflow, power and loads of the downstream turbine.

4.1. Inflow

At first, we evaluate the statistics of the ambient flow along the met mast. Fig. 4 compares the results of the simulations to the measurements of the sonics as well as the provided inflow turbulence box.

The TKE as well as the individual velocity variances in the DWM simulations are in close agreement with the imposed turbulence box. This can generally be expected because the imposed turbulence in DWM models is passively advected through the domain and only evolves if being superimposed by occurring wakes. The observed deviations are thus exclusively related to interpolation errors or predefined scaling factors in the respective framework. The mean velocity profiles of the LES cases somewhat closely match with the measured velocities. The most notable deviation is slightly elevated velocities in the LESDTU case near the ground. These are likely to be attributed to the applied symmetry boundary condition at the bottom. In this case, the imposed turbulent velocity field is not subjected to any retarding wall shear stress while propagating downstream. The TKE of the LES cases, on the other hand, only falls within the bounds of the measurement uncertainty at hub height. In all cases, regardless of the bottom boundary condition, the imposed near-wall fluctuations cannot be sustained between the inlet and the met mast location. As the imposed turbulence is typically not in equilibrium with the LES solution, a continuous decay with downstream distance can generally be expected, see, e.g. Refs. [39,86]. Furthermore, the different discrepancies in the individual variances indicate that the decay is highly anisotropic while also differing substantially between the models. The precursor reference case notably better matches the stream-wise variance while showing similar deviations in the lateral and vertical component.

4.2. Power and thrust of the upstream turbine

The mean mechanical power, \( \overline{P} \), and total thrust force, \( \overline{T} \), acting on the rotor of WT1 as well as the corresponding standard deviations are compared in Fig. 5. The means are plotted against the averaged mean stream-wise velocity at the locations of the three sonics (\( \overline{\nu_z} \)), providing an estimate of the mean inflow across the rotor swept area. The standard deviations are plotted against the corresponding averaged stream-wise velocity variance (\( \overline{\nu^2} \)).

In all simulations, the mean power lies below the measured values derived from the SCADA data, ranging from \(-0.8\%\) for LESUU to \(-18.3\%\) for DWMNREL. The mean power and thrust of the simulations appear weakly correlated to the mean inflow velocity, potentially explaining some of the differences among the models and the measurement. Note that the thrust is not measured in the experiment. A slightly stronger correlation can be found for the standard deviation of the power (and thrust) and \( \overline{\nu^2} \).

4.3. Wake inflow of the downstream turbine

Vertical and horizontal characteristics of the wake of WT1 are depicted in Fig. 6. As mentioned in Sec. 2, no direct freestream measurement of the wake flow is available for the discussed cases. The only available references are the wind speed measurement of the nacelle cup anemometer as well as the velocity measurements of the Pitot tube mounted on the blade of WT2. Comparing upstream velocities with the former can entail large uncertainties because of the position of the anemometer behind the rotor [87,88]. As for the latter, the measurement position is potentially more suitable. Still, induction effects due to the vicinity of the blade cannot be ruled out. Furthermore, the measurements for a fixed spatial position can only be estimated from finite size bins of the azimuth angle \( \phi \), due to the rotation of the blade. For the data points shown in Fig. 6, statistics were obtained for bins of \( \Delta \phi = 12^\circ \) around the target azimuth positions of 90° and 270° for the horizontal

![Fig. 4. Mean stream-wise velocity \( \overline{\nu} \) (a), resolved TKE (b) as well as the resolved stream-wise (c), lateral (d) and vertical velocity variances (e) along the vertical at the met mast location for Case A. All quantities are normalised by the mean stream-wise velocity of the sonic \( \overline{\nu_z} = \overline{\nu}_0 \), hereafter denoted as \( \nu_0 \). The error bars show the uncertainty estimated via the integral length scale of the involved velocity components using the method proposed by Lenschow et al. [85]. Note that the steady-state RANS simulation (RANSUU) only provides modelled (sub-grid scale) turbulent quantities that are not compared here for the sake of brevity.](image-url)
profiles and 0° and 180° for the vertical profiles. The effective measurement time for the shown statistics therefore only amounts to a fraction of the bin length, explaining the significantly larger uncertainty than for the cup anemometer measurements.

The lateral offset of the wake center with respect to the rotor center of WT2 grows with increasing downstream distance due to the orientation of the line samples normal to the x-axis (see Fig. 3). Eventually we find a half-wake in flow for WT2 in all simulations.

Four diameters upstream of WT2, all models except RANSUU exhibit a typical near-wake velocity profile. A clear correlation between $T$ and the magnitude of the velocity deficit cannot be observed. Other aspects like the smearing width in the ALMs or more fundamental numerical differences between the models thus seem to dominate the discrepancies in the flow in this region. At $x = 2D$, the velocity profile in all simulations reaches a Gaussian far-wake state. Hence, the downstream distance of the laminar-turbulent transition of the wake of WT1 is similarly predicted by all LES and DWM models.

Generally, the wake recovery in the RANS case remains faster than in the other models, indicating a higher numerical and/or sub-grid scale dissipation. One diameter upstream of WT2, the LES and DWMNREL are in quite close agreement in terms of the velocity deficit. DWMDTU and RANSUU predict a somewhat larger and smaller velocity deficit, respectively. For the DWMDTU model, the steeper deficit would be reduced using a slightly higher value of the parameter $k_{amb}$ linking the eddy viscosity to the ambient turbulence intensity. The value of 0.10 used here, provided a good correlation with measured loads in the Egmond Aan Zee wind farm [59]. Likewise, a higher value of $k_{amb}$ reduces the wake-generated turbulence from meandering, which also would give an improved correlation in the present benchmark for the DWMDTU model.

Furthermore, it can be appreciated that the models lie close to the mean of the nacelle anemometer measurement as well as the Pitot tube data point at $\varphi = 270°$ (positive y-direction). At $\varphi = 90°$ (negative y-direction), the Pitot tube depicts a considerably larger velocity deficit. A similar picture is shown for the vertical velocity profiles. Here, the Pitot tube shows lower mean values both above and below hub height while the nacelle anemometer agrees better...
with most simulations. Larger deviations between the models are found in terms of \( \overline{u^3} \). Particularly in the near-wake DWMNDTU for instance, shows distinctly larger peaks than most LES models. However, the wake-added turbulence model in DWMNDTU has been calibrated for downstream distances larger than 1 – 2 D. On the other hand, DWMNREL underpredicts \( \overline{u^3} \) because DWMNREL does not include a wake-added turbulence model as of yet. Generally, it can be noted that the flow downstream of WT1 seems to be dominated by wake-generated turbulence. After all, the aforementioned differences in the inflow turbulence are not reflected in corresponding differences in the wake. With increasing down-stream distance, both magnitude and shape of the \( \overline{u^3} \) profiles converge. At \(-1D\) the differences between the models are significantly smaller than further upstream, in particular, between the LES models. Moreover, they rather closely match with the measured \( \overline{u^3} \) of the Pitot tube.

In Fig. 7a, we therefore subtract the mean of the blade normal velocity, \( \overline{u_0} \), from each simulation from the mean obtained for each azimuthal bin, \( \overline{u_0} \). This way, we focus on the comparison of the wake shape while reducing uncertainties related to the aforementioned aspects. Also note that \( u_0 \) and \( u \) shown in Fig. 6 do not exactly comply because the wind direction is not aligned with the x-axis in this case.

In the measurements, the maximal velocity lies at \( \phi = 72° \). With increasing azimuth angle, \( \pi_\alpha \) decreases because the blade is moving to lower heights and toward the core of the wake of WT1. The minimum is then found at \( \phi = 240° \). The azimuthal course of \( \overline{u^3} \) is generally well-captured by the LES and RANS models. This refers to both the magnitude and phase of the azimuthal oscillation. Smaller local deviations can be observed over certain intervals of \( \phi \). Though not largely off, more notable deviations are found for the DWM models featuring a larger magnitude of the oscillation and a positive shift in the phase. The characteristic shape of \( \overline{u^3}(\phi) \) of the simulations shows more distinct deviations from the measurements. This applies particularly to the peak of the velocity variance in the measurements at \( \phi = 108° \). Still, for \( \phi > 200° \) three of the LES models are in closer agreement with the measurement. DWMNREL is lower than the other models due to the lack of a wake-added turbulence model.

### 4.4. Power and loads of the downstream turbine

The mean and standard deviation of the power and thrust of WT2 are contrasted in Fig. 8. As opposed to Fig. 5, we now plot against the mean and variance of \( u_{\alpha} \), respectively, as it states the most suitable inflow velocity data available from the simulations and the measurements for WT2.

The simulation data show a significantly better correlation between \( P \) and \( \pi_\alpha \) than the one shown in Fig. 5. The overprediction of the power in the majority of the models consolidates that the velocity deficit is indeed underpredicted, as indicated by the results in Fig. 6. A similarly good correlation is found between \( \sigma(P) \) and \( \sigma(T) \) and \( \overline{u_\alpha u_{\alpha}} \). Furthermore, the SCADA data more closely agree with the trend emerging from the simulation data. This again corroborates the existence of a systematic offset in the mean of the blade normal velocity, as discussed for Fig. 7. The mean forces along the blade of WT2 are compared against the force measurements from the pressure taps in Fig. 9.

In the mid third of the blade, the closest match with the tangential force measurements is found for DWMNREL and LESUU,LLBM. Toward the tip, all LES and RANS models except LESDTU overestimate \( F_n \), LESDTU generally predicts a smoother decline of the force toward the tip and thus better captures the overall trend of the measurements. Referring to the discussions in Ref. [77], the latter can clearly be attributed to the utilized tuning-free tip correction model. Note that the other LES models used an empirical tip-correction model or no correction at all (see Tab. 5). Moreover, it should be noted that the force measurements do not capture viscous forces. In contrast, all turbine models rely on tabulated lift and drag values that do include them. Bangga and Lutz [32] therefore argue that a closer agreement in \( F_n \) not necessarily implies a better model prediction. On the other hand, the differences in \( F_n \) are qualitatively consistent with the differences in the power, indicating that the neglected component cannot have a major contribution. The relative differences between the models in the normal forces, which are significantly larger in magnitude than the tangential forces, are considerably smaller. At the same time, all models consistently overestimate \( F_n \) when compared to the measurements. As for \( F_n \), a direct comparison to the measurements is
more straightforward because the contribution of the viscous component is deemed negligible. In addition, the lower measured forces are consistent with the higher velocity deficit demonstrated by the nacelle anemometer and Pitot tube. Nonetheless, it should be mentioned that selected cases of the DanAero experiment investigated within Task 29 showed inconsistencies when comparing measurements of the pressure forces and bending moments; see Ref. [33]. A small bias in the measured $F_n$ cannot be fully ruled out.

In Fig. 10, we compare the azimuthal variation of the mean normal force and the corresponding standard deviation. Analogously to $u_n$, the force measurements exhibit a maximum at the right-hand side of the rotor plane (seen from upstream) at about $90^\circ$ azimuth. The minimum is found at $45^\circ$ to $270^\circ$. The phase of this azimuthal oscillation of $F_n$ is quite closely captured by the majority of the models. In contrast, the agreement in terms of the magnitudes of the minima and maxima differs greatly depending on the model and blade section. Expectantly, we find a strong correlation between the match in terms of the force and the match in terms of the inflow velocity in the middle of the blade. In addition to the direct comparison of the velocity, this provides further indications of how well the wake is captured by the different models. Closer to the tip, such conclusions become less straightforward, as the magnitude of the force is no longer dominated by the inflow velocity and increasingly affected by the specific choices made in the turbine model setup (such as the tip correction or, e.g., the smearing width in the ALMs). In contrast to the mean, the standard deviation of the measured normal forces does not follow a similarly distinct sinusoidal curve. Moreover, it generally exhibits less azimuthal variation. The trends exhibited by the simulations differ more significantly from the experimental data in large parts. Furthermore, the fluctuations are consistently underpredicted, which appears contradictory to the comparison of the velocity variances (see Figs. 6 and 7). Unfortunately, the data do not allow for a conclusive explanation of the issue other than the aforementioned measurement uncertainties. Lastly, we evaluate the power spectral density (PSD) of the normal force, see Fig. 11.

For all cases, the most energy is found at the 1P frequency. Throughout the simulations, this peak is consistently more distinct than in the measurements including a larger magnitude. At $r/R = 0.925$, the simulations also show clear secondary peaks at the higher harmonics, which is not the case for the measurement. In comparison to the PSD in Case B (shown later in Fig. 19), these deviations seem to originate from varying $\omega$ in the measurements that is not replicated in the simulations.

5. Case B - full wake

Following the same structure as before, Case B featuring a full wake inflow for WT2 will be discussed in the following.

5.1. Inflow

Fig. 12 depicts the inflow statistics along the met mast for Case B. Again, the DWM models show the closest agreement with the statistics of the imposed turbulence box along the entire met mast. In the LES cases, the mean velocity is exaggerated by 2–5%. The largest deviations are again found at the lowest measurement point. Somewhat similar characteristics are also found for the TKE and the individual velocity variances. When compared to Case A, the largest difference is a more notable underprediction of the
vertical variance across all LES models. Because $w_0$ is generally larger than in the previous case, this possibly relates to the stronger stratification that is not captured by the simulations.

5.2. Power and loads of the upstream turbine

The mean and standard deviation of the power and thrust, respectively, of WT1 are compared in Fig. 13. As for $P$, the agreement with the simulations ranges from 1.2% overprediction for LESUU to 16.9% underprediction for DWMDTU. Generally, both $P$ and $T$ are similarly well-correlated to the mean inflow as in Case A. Equally similar is the higher power-to-velocity ratio of the SCADA data when compared to the trend in the simulations. Hence, overpredictions of the inflow velocity tend to imply a better match with the measured power than cases with a close match in the former. Surprisingly, no clear trend in the standard deviations can be observed for this case.

5.3. Wake in inflow of the downstream turbine

The horizontal and vertical wake characteristics are shown in Fig. 14.

In terms of the mean velocity, the simulations compare similarly as in Case A; while more pronounced differences can be observed in the near-wake, the far-wake velocity profiles are quite similar in terms of velocity deficit and wake expansion. On the other hand, the agreement with the Pitot tube and nacelle anemometer measurements is slightly worse. Particularly, the former shows a noticeably larger wake deficit on the right of the rotor center (negative $y$). The wake in the measurements thus seems to exhibit an asymmetry along the x-axis that is not captured in the simulations. The origin of the asymmetry can be the vertical wind veer, as well as the rotation of the wake. Unfortunately, it is not straightforward to determine from the available data whether...
the asymmetry originates from an overall wake deflection or an asymmetric wake recovery. When compared to Case A, we observe a notably larger spread in the modelled stream-wise variance. The closest match with the Pitot tube measurements is found for the LES models. DWMDTU quite severely overpredicts the variance. Fig. 15 compares the statistics of the blade-normal inflow velocity.

The azimuthal variation of $u_{in}$ complements the picture of the deviations in the wake between the simulations and measurements discussed above. Firstly, the shear in the measured wake inflow appears to be significantly larger, as indicated by a larger difference between the lowest ($\varphi = 180^\circ$) and the highest ($\varphi = 0^\circ$) position of the blade. Secondly, the measured wake profile is clearly shifted with respect to the simulations with a minimum at $\varphi \approx 150^\circ$. The velocity variance in the simulations shows no strong variation with the azimuth angle while mainly differing in the magnitude. Conversely, the measured $u_{in}$ is clearly sinusoidal with a maximum to the right of the rotor center.

5.4. Power and loads of the downstream turbine

Statistics of the power and thrust of the downstream turbine are provided in Fig. 16. The majority of the models overpredict the mean power of WT2. As opposed to Case A, the measured power more closely aligns with the power-velocity trend given by the simulations. With the exception of LESUU, the simulations and the measurement also depict a common trend in terms of the dependency of $\sigma(P)$ and $\sigma(T)$. In comparison, such a clear trend cannot be observed for the mean and standard deviation of the thrust. Fig. 17 depicts the mean tangential and normal forces along the blade.

Between the inner two blade sections, all models predict a relatively constant tangential force with the magnitude lying below the measurements. Toward the tip, characteristic differences between the models increase. Of the LES and RANS models, again only LESDTU predicts a decay of $F_t$. The other models show a typical bump that originates from the spurious induction of the trailing tip vortex \[77\] (or the vortex sheet for the ADM in RANSUU, respectively \[91\]). When compared to Case A, DWM_DTU is in closer agreement with most of the LES models. The normal force in the simulations is again consistently higher than in the measurements. This also agrees with the larger velocity deficit found in the experiment. The azimuthal variation of the normal force and its standard deviation are compared in Fig. 18. The mean measured force follows a clear sinusoidal curve. Depending on the radial position and result, the minimum is found between $160^\circ$ and $170^\circ$ azimuth in most models. In line with the velocity measurements, this indicates a slight asymmetry in the wake. In comparison to Case A, the models differ more severely in terms of the overall azimuthal characteristic of $F_n$. For example, LESUU and RANSUU predict a sinusoidal variation that is similar to the measurements. Yet, the minimum in these cases is closer to $190^\circ$. On the other hand, LESDTU and LESUU,LBM exhibit a comparably small variation with the azimuth angle, mainly due to differences in the vertical profile of the inflow velocity (see Fig. 14). DWMDNUCL
predicts an azimuthal variation that is opposite to the other results, but also found that the azimuth variation depended on the time window analyzed, implying that these results might not be statistically converged to some extent. The standard deviation of the measured normal force exhibits minima at similar azimuth angles as the mean, yet the curve is less sinusoidal. While none of the

Fig. 14. Wake characteristics upstream of WT2 in Case B. For further details, see of Fig. 6.

Fig. 15. Azimuthal course of blade inflow velocity statistics at r/R = 0.51 of WT2 in Case B. For further details, see of Fig. 7.

Fig. 16. Mean power, $P$ (a), and mean total thrust force, $T$ (c), of WT2 vs. mean of the stream-wise blade velocity $u_n$ for Case B. Corresponding standard deviation, $\sigma(P)$ (b), and $\sigma(T)$ (d), vs. the stream-wise velocity variance, $u_n^2$. 

H. Asmuth, G.P. Navarro Diaz, H.A. Madsen et al. Renewable Energy 191 (2022) 868–887
models consistently predicts this characteristic at all radial stations. LESUU, LESUU,LBM, LESUU and LESDTU generally show the closest fit with the measurements. In Fig. 19, we compare the PSD of the normal force at two of the blade sections.

When compared to Case A, the peak in the measurements at the 1P frequency is more pronounced. Moreover, the force shows more distinct peaks at the higher harmonics of the rotor frequency, especially at the outer blade section. The spectra of the modelled forces are in close agreement with the measurements. This corroborates the conjecture that some of the differences found in Case A do originate from deviations in the rotational speed.

6. Discussion

As a basis for a final discussion of the results, we illustrate the $L^2$-relative error of various discussed quantities in Table 3, where

$$L^2 = \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}Y_i^2} \cdot 100$$

with $Y_i$ being a measured value, $X_i$ being the respective modelled value and $n$ being the number of data points. The errors of the mean velocity and TKE along the met mast ($\bar{V}_{MM}$ and $\overline{TKE_{MM}}$) are evaluated at the heights of the three sonic anemometers. The errors of the normal force refer to the average relative error of the corresponding quantity at the four radial blade sections. Generally, the provided error values shall facilitate a summarizing quantitative comparison with the measurements. However, we would like to emphasize that all given values are eventually affected by the various uncertainties mentioned throughout Sects. 4 and 5. This refers to the uncertainties of the measurements themselves, potential comparability issues, such as the lacking viscous force component in the measurements, and, generally, the rather short length of the compared data of only 10 min per case.

6.1. Model evaluation

The most informative quantity for the evaluation of the model performance in this study is the error in the azimuthal variation of the normal force. For the majority of the models this error is found in the range of 15–20%. Bangga and Lutz [32] compared BEM and fully resolved rotor simulations of DanAero cases in unwaked inflow conditions. The average relative errors of $F_{n}$ reported for the four blade sections lie in the range of 3–9%. Grinderslev et al. [30] compared aeroelastic fully resolved rotor simulations of high yaw angles, unwaked inflow cases against DanAero measurements with corresponding errors of 5–7%.$^1$ In comparison, the results presented in this benchmark can arguably be appreciated given the higher complexity associated with the modelling of the wake inflow.

The main motivation for this study is the plain comparison of the current state of the investigated simulation frameworks. Beyond that, it is obviously desirable to pinpoint explicit model weaknesses to target future model improvement efforts. Some aspects have already been discussed in Sects. 4 and 5. In the following we shall summarize some of the main findings, deficiencies and potential remedies for the three model families.

6.1.1. RANS

- The most striking deficit is the overly strong decay of the wake which we mainly attribute to the choice of turbulence model. It generally illustrates the difficulty in finding RANS parametrizations that are suitable for both boundary layer and wake flow regimes [92]. Potential remedies might be found in other RANS parametrization, as proposed in Refs. [92–94].

6.1.2. DWM

- Sufficient spatial coherence of the imposed inflow in all three directions is crucial to evoke a realistic wake meandering and, consequently, achieve realistic wake turbulence statistics.
- Cases with low inflow turbulence remain challenging because the only driving factor for the wake meandering considered in the model (large-scale structures in the ambient flow) is weak.
- Additional parametrizations of wake-added turbulence can be a remedy for the lacking meandering in low ambient turbulence but are difficult to tune (see, e.g., DWM$_{DTU}$ in Fig. 14).

6.1.3. LES

- All LES cases are quite consistent in terms of the predicted wake statistics, despite the different numerical approaches and SGS models.

$^1$ Note that the explicit errors are not given in the original paper but were kindly provided by the authors upon request.
The higher spatial resolution in the wake region in LESDTU ($\Delta x = D/60$) does not seem to pay off in terms of a higher accuracy (of the compared statistics), when compared to the other LES cases. Statistics up to second-order thus seem to be sufficiently captured with lower spatial resolutions of $D_x = D/32$ (as in LESUU and LESUU,LBM), in line with other studies [95,96].

The generality of tuning-free tip corrections [77] is of paramount importance for accurate load distributions in actuator models.

The biggest challenge for LES remains the generation of turbulent inflow data that matches the measured flow conditions at some reference point in the domain (as further illustrated in Sec. 6.2).

Beyond the aspects mentioned above, further conclusions remain rather speculative due to the limited knowledge about the events upstream of WT2. This mostly relates to the lacking freestream measurements of the wake inflow. But also time-resolved SCADA data and load measurements of the upstream turbine or additional upstream velocity measurements would naturally help to narrow down the sources of the errors discussed above. After all, for a wake inflow case, errors in the modelling of the response of the downstream turbine can originate from the turbine model itself, the prescribed inflow, the modelling of the upstream turbine and its wake or, most likely, a combination of the four. One example is that even low errors in the inflow or power of WT1 do not necessarily correlate with low errors in the power or loads of WT2. More generally, this rather low correlation between different errors is illustrated by the large scatter shown in Tab 3.

### 6.2. Inflow generation

The reproduction of a measured inflow in time-resolved fluid dynamic simulations still presents a major modelling challenge, not only in the context of wind turbine simulations [97]. Regardless of
the method used to generate the inflow turbulence, for LES, the problem remains that imposed synthetic turbulence evolves as it propagates downstream. A decay of TKE after the inlet and even local changes in the time-averaged velocity downstream are therefore inevitable.

With regards to this study, the issue is illustrated in Fig. 20 by comparing the downstream evolution of the TKE of LESUU and LESUU,pc.

For both cases, a significant decay of TKE can be noticed between the inlet and WT2 along the undisturbed line. In Case A, the magnitude of the TKE drops by 14% between the inlet and the met mast position. Only downstream of the met mast the decay vanishes, yet the magnitude remains above the measured TKE. In Case B, the TKE keeps decaying up until WT2 with a decrease of 32% between the inlet and the met mast. On the other hand, the TKE with the precursor inflow remains nearly constant throughout the domain. The effects of the two inflow methods on the wake can be observed along the line passing through the rotor center. In both cases, the laminar-turbulent transition of the wake of WT1 seems to occur earlier in LESUU, indicated by an earlier increase in TKE. The example illustrates that the use of synthetic inflow turbulence in LES for this particular case is arguably not ideal. Further calibration runs might yield better scaling factors to match the statistics at the met mast. Furthermore, the anisotropy of the decay, observed in Figs. 4 and 12, seems to require individual scaling factors for each velocity component. Still, both measures do not remedy the underlying problem, i.e. the non-equilibrium of the imposed inflow with the LES solution downstream. Expectantly, this problem is not evident in the precursor case. On the other hand, even this case fails to adequately reproduce the turbulence statistics at the met mast, possibly due to neglecting stratification and/or insufficient tuning of the driving pressure gradient and surface roughness.

7. Conclusion

The need for reliable modelling approaches for wind turbines and farms has led to numerous collaborative experimental validation studies over the past two decades. Nevertheless, validations against full-scale measurements of detailed blade loads in waked inflow conditions remain rare. At the same time, they represent one of the most stringent tests of the modelling capabilities of a framework. Eventually, they scrutinize the comprehensive accuracy in simulating the interaction of the ABL, wakes of upstream turbines and the aerodynamic response of the turbine of interest. In this study, we presented a comprehensive validation of six model frameworks of different fidelities against measurements of a 2.3 MW wind turbine in waked inflow conditions. The models were evaluated in terms of the ambient flow conditions, the wake flow, power and thrust of the upstream and downstream turbine and, ultimately, the loads of the downstream turbine.

The two benchmark cases revealed that the majority of the models are able to capture the main characteristic features of the wake inflow as well as the resulting forces acting on the downstream turbine. This result is generally encouraging, especially in view of the simplifications in the case setup (neglecting stratification, constant rotational speed and yaw). Despite the explicit error values given in the study, we refrain from a summarizing quantitative ranking of the performance of the models due to the discussed uncertainties in the measurements. Nonetheless, it can be concluded that the compared quantities were not consistently better captured by the models with the highest fidelity, i.e. LES. This

| Table 3 | Comparison of the L²-relative error of a selection of the discussed quantities of Case A (a) Case B (b). Dark red corresponds to the highest error in the respective quantity, dark blue to the lowest. |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|
|         | (a)     | (b)     |         |         |         |         |         |         |
| η [%]   | 2.9     | 1.1     | 1.5     | 2.2     | 4.6     | 3.9     | 2.6     | 3.2     | 2.4     | 1.2     | 5.4     | 3.9     | 2.7     | 2.8     |
| TKE [m²/s²] | 12.0   | 15.4   | 10.0   | 20.4   | 21.1   | 16.0   | 15.3   | 16.8   | 21.4   | 16.0   | 23.5   | 17.1   | 18.9   | 17.7   |
| F₄ F₄   | 47.2   | 67.2   | 62.6   | 28.1   | 26.9   | 48.3   | 39.9   | 52.2   | 173.6  | 56.9   | 68.9   | 107.0  | 73.1   | 80.4   |
| σ F₄     | 34.0   | 30.9   | 25.5   | 39.8   | 26.7   | 32.8   | 35.4   | 17.8   | 53.2   | 17.1   | 20.3   | 12.1   | 12.6   |
| F₃ [%]   | 1.9    | 16.9   | 15.9   | 11.3   | 5.1    | 0.8    | 5.5    | 1.6    | 17.1   | 16.9   | 5.5    | 7.1    | 1.2    | 1.7    |
| F₂ [%]   | 27.4   | 5.5    | 15.4   | 7.8    | 19.6   | 24.7   | 26.1   | 40.8   | 4.5    | 3.0    | 3.8    | 7.2    | 9.3    | 23.8   |

Fig. 20. Comparison of the downstream evolution of the TKE at hub height in the rotor center (y = 0 D) and at y = −2.5 D for LESUU and LESUU,pc in Case A (top) and B (bottom). The turbine positions are marked by dashed grey lines. The black circle shows the TKE of the sonic anemometer at hub height with the error bar marking the measurement uncertainty [85].
refers to both mean and second-order quantities. To some extent, we can relate this to the difficulties in matching the ambient flow conditions at the met mast with LES. Similar problems have been reported in previous studies and can be stated as a persistent challenge for the method [12].

The entirety of measurement comparisons in this benchmark clearly illustrate that validations of wind turbine models should not be limited to integrated quantities such as power or bending moments. Such quantities can hide crucial local deficiencies of the model frameworks either in terms of the wake flow or the modelling of the blade forces. In addition, over- and underestimation can compensate for each other and remain undetected. Similar misjudgments can arise from comparisons restricted to time-averaged quantities like the mean forces discussed in this study. Furthermore, the study shows that precise statements about the origin of modelling errors and potentials for model improvements can only be made with sufficient knowledge about all relevant upstream events.

In line with the recent study by Doubrawa et al. [12], it can be emphasized that future full-scale measurement campaigns should aim to reduce the instrumentation limitations of past experiments. Regardless of the flow scenario, this implies more detailed measurements of the undisturbed ambient flow. Particularly, this refers to measurements that go beyond the turbine height, allowing for a quantification of a larger extent of theABL, and that are not limited to a single met-mast location. This inevitably also has to include temperature measurements such that a quantification of the atmospheric stability is possible. Several met-masts, remote sensing instruments like scanning LIDARS or a combination of the two can be a suitable approach. Additional soundings of the boundary layer height and inversion strength (e.g., using automatic LIDARS and ceilometers [98,99]) can be useful indicators, particularly to set-up boundary and initial conditions in LES. Similar requirements apply to the measurements of wakes. This study again highlights that a meaningful evaluation of a modelled wake requires multiple measurement locations in all three spatial directions, ideally including measurements of all three velocity components. After all, the rotation, decay and deflection of the wake are all closely interlinked. Any missing aspect about the downstream evolution of the wake thus inevitably makes the search for model deficiencies more speculative. Nacelle-mounted LIDARS appear to be one viable solution for this task, despite persistent challenges such as probe volume effects and the limitation to the line-of-site velocity component [12,100,101]. The comparatively detailed blade load measurements conducted in the DanAero experiment (i.e., multiple local pressure measurements and strain gauges) clearly prove to be a valuable asset for the evaluation of model performance. Yet, only the combination of inflow and load measurement allows for a clear assessment of actuator or BEM-based turbine models in waked inflow conditions. We can therefore highly recommend similar set-ups for future measurements campaigns even though expensive and challenging to calibrate [33]. Further valuable additions can be strain gauges on the tower, enabling reference estimations of the overall thrust force. Lastly, it should be stressed that changes of the operational mode of the measured turbine (like fixing the pitch and rotational speed of WT2 in Case A) can be a simple, yet very useful measure to reduce uncertainties and facilitate model comparisons.

Despite the limitations in the comparison to the DanAero experiment, the discussed cases remain a valuable database for future studies. Interesting aspects for further investigations would be simulations with less simplifications than in this work, e.g. including a torque and pitch controller. Furthermore, with more tuning, a better match of the inflow can surely be achieved using LES, reducing the uncertainties related to this aspect.

CRediT authorship contribution statement

Henrik Asmuth: Conceptualization, Software, Data curation, Formal analysis, Writing — original draft, (lead), Project administration, ren. Gonzalo P. Navarro Diaz: Software, Writing — original draft, (supporting). Helge Aagaard Madsen: Software, Writing — original draft, (supporting), Funding acquisition. Emmanuel Branlard: Software, Data curation, Writing — original draft, (supporting). Karl Nilsson: Conceptualization, Data curation, Project administration. Jason Jonkman: Writing — original draft, (supporting), Funding acquisition. Stefan Ivanell: Conceptualization, Funding acquisition, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study emerged from activities organised within IEA Task 29 Phase IV. The authors would like to thank IEA Wind for facilitating the project. The participation of Uppsala Universitett in IEA Task 29 was financed by the Swedish Energy Agency. Most simulations were run on computer resources provided by the Swedish National Infrastructure for Computing (SNIC) at the National Supercomputing Center (NSC) under the 927 Project No. SNIC 2020/1-10. Their support is gratefully acknowledged. The DTU contributions have been carried out based on funding from EUDP 2018-1, contract J.nr. 64018-0084: “Participation in IEA Task 29: Full Scale Wind Turbine Aerodynamics, -elasticity and -acoustics”.

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

A. Numerical Set-ups

A summary of the details of numerical setups of the LES and RANS simulations is provided in Tab. 4. Further information about the wind turbine models is given in Tab. 5.
Table 4
Numerical configurations of the RANS and LES models: overall dimensions of the computational domain $L_{cha}^{e}$, dimensions of the finest (isotropic) grid region in the turbine and wake vicinity $L_{cha}^{e}$, with corresponding grid spacing $\Delta x = \Delta y z$; total number of grid points (cells); utilised bottom boundary conditions (stress BC implying the use of a wall function with corresponding aerodynamic roughness $\kappa$) and time step $\Delta t$.

| $L_{cha}^{e}$ [m] | $L_{cha}^{e}$ [m] | $N_{x}$ | $N_{y}$ | $N_{z}$ | $\Delta x$ | $\Delta y$ | $\Delta z$ | $N_{points}$ | $bc$ | $\Delta t$ [s] |
|------------------|------------------|-------|-------|-------|----------|----------|----------|-----------|------|----------|
| 20.0 x 10.0 x 10.0 | 16.6 x 10.0 x 18.75 | 16.6 x 10.0 x 12.5 | 37.5 x 10.0 x 12.5 | 17.0 x 7.5 x 12.5 | 1.5/2 | 1.1/10/1/20 | 1.1/2 | 1/10/1/20 | 1.5/2 | 1.1/10/1/20 |
| 12 x 4 x 4 | 16.6 x 10.0 x 18.75 | 16.0 x 10.0 x 19 | 11.2 x 5.0 x 14 | 1.5/2 | 1.1/10/1/20 | 1.1/2 | 1/10/1/20 | 1.5/2 | 1.1/10/1/20 |
| 1/32 | 1/10/1/20 | 1.1/2 | 1/10/1/20 | 1.5/2 | 1.1/10/1/20 |

Table 5
Wind turbine models configurations: Radial spacing of blade elements $\Delta s$, smearing width $\epsilon$ and utilised tip correction model.

| Model | $\Delta s/\epsilon$ | Tip correction |
|-------|------------------|----------------|
| LES/LBM ALM | 2 | 1.5 | – |
| RANS/LBM | 2 | 2 | Shen [50,51] |
| LES/LBM/LES/LBM | 2 | 2 | Shen [50,51] |
| DWMFL BEM | 1/10 | 2.5 Meyer Forsting et al. [77-79] |
| Prandtl [102] |
| DWMRE BEM | 0.6–1/5 | 1/32 1/10, 1/20 1/32, 1/32 1/60 |

References

[1] P. Veers, K. Dykes, E. Lantz, S. Barth, C.L. Bottasso, O. Carlson, A. Clifton, J. Green, P. Gren, H. Hofttiken, D. Laird, V. Lehtinen, J.K. Lundquist, J. Manwell, M. Marquis, C. Meneveau, P. Mioriart, X. Munduate, M. Muskulus, J. Naughton, L. Pao, J. Paquette, J. Peinke, A. Robertson, J. Sanz-Rodrigo, A.M. Sempreviva, J.C. Smith, A. Tuohy, R. Wiser, Grand challenges in the science of wind energy, Science 366 (6464) (2019), https://doi.org/10.1126/science.aar5150.

[2] C. Meneveau, Big wind power: seven questions for turbulence research, J. Turbul. 20 (1) (2019) 2–20, https://doi.org/10.1080/14685248.2018.1544664.

[3] F. Porté-Agel, M. Bastankhah, S. Shamsoddin, Wind-turbine and wind-farm flows: a review, Boundary-Layer Meteorol. 174 (1) (2020) 1–59, https://doi.org/10.1007/s10546-019-00563-0.

[4] H. Aas, H. Olivier-Espinosa, S. Ivanell, Actuator line simulations of wind turbine wakes, J. Phys. Conf. Ser. 524 (2014), 012145, https://doi.org/10.1088/1742-6596/524/1/012145.

[5] H. Aas, H. Olivier-Espinosa, S. Ivanell, Assessment of turbulence modelling in the wake of an actuator disk with a decaying turbulence inflow, Appl. Sci. 9 (8) (2019), https://doi.org/10.3390/app9081530.

[6] H. Asmuth, H. Olivier-Espinosa, S. Ivanell, Assessment of turbulence modelling in the wake of an actuator disk with a decaying turbulence inflow, Appl. Sci. 9 (8) (2019), https://doi.org/10.3390/app9081530.
