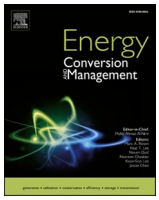
Energy Conversion and Management 296 (2023) 117708

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman



Wake and power prediction of horizontal-axis wind farm under yaw-controlled conditions with machine learning

M.E. Nakhchi *, S. Win Naung, M. Rahmati

Faculty of Engineering and Environment, Northumbria University, Newcastle upon Tyne NE1 8ST, UK

ARTICLE INFO

Keywords:
Wind farm
XGBoost
Yaw control
Power estimation
Machine learning

ABSTRACT

The main objective of this study is to employ the Extreme Gradient Boosting (XGBoost) machine learning algorithm to predict the power, wake, and turbulent characteristics of horizontal-axis wind farms under yaw-controlled conditions. For this purpose, a series of high-fidelity numerical simulations using LES method are performed over tandem NREL-5 MW wind turbines to generate the input data for training and testing in machine learning analysis. It is observed that XGBoost is more accurate for wake prediction of the yaw-controlled wind farms compared to ANN, which was used in previous studies. The results illustrate that XGBoost can predict the power with a mean deviation of 0.94 % for different yaw angles, while ANN can estimate the power generation with a mean deviation of 2.15 % for various tested yaw angles. At far wake regions ($X > 2000$ m) of the second wind turbine, the deviations reach below 1 %. Moreover, XGBoost requires a much shorter training time, 87.5 % faster than ANN. The power production of both wind turbines can be predicted more accurately with XGBoost compared to ANN. The wake prediction time of XGBoost is just 0.105 sec, while this time is 4.480 for the ANN model. In conclusion, XGBoost provides a significant reduction in error and training time compared to ANN and deep learning algorithms over yaw-misaligned wind farms.

1. Introduction

In order to reduce the total cost of wind power generation, wind turbines are usually clustered to form a wind farm. However, if the downstream turbines run in the wake of the upstream ones, their performance may be significantly affected. Wind turbines operating in the wake regions typically produce less power and will face severe structural stresses than those running in freestream flows. Wind farm layouts generally are optimised to minimise the effects of the unfavourable wake. Several wake control techniques, such as yaw control, tilt - control, and cone angle control techniques, are suggested to deflect the wake of upper-layer wind turbines in complex wind farm layouts [1].

One of the most common approaches to evaluate the wind-turbines aerodynamic performance and wake profile is to use computational fluid dynamics (CFD), for example large-eddy simulation (LES), which provides a more precise characterisation of flow structure over wind turbine blades than the Reynolds averaged Navier Stokes (RANS) model [2–4]. The improved mesh used to determine the boundary layer makes CFD modelling of a wind farm with tens of turbines highly computationally costly. The actuator-disk method (ADM) and actuator line method (ALM) coupled with large eddy simulation (LES) can simulate complex wind turbines by replacing the blades with actuator points [5]. Although the computational cost of LES is still quite expensive to resolve the turbulent characteristics in wake regions, the ALM-LES method has been successfully implemented to predict the wind turbine wake [6,7]. To explore the time-averaged wake structures, the RANS method is more computationally efficient than the LES. The common - turbulence models in RANS, like standard k-ε, cannot accurately compute the turbulent flow in the turbine wake under atmospheric boundary layer (ABL) conditions, resulting in a failure in wake predicting, as noted by [8,9]. In general, low-fidelity CFD modelling can estimate wake characteristics that are more specific and precise. Therefore, high-fidelity CFD models may be used for various wind turbine wake structures, power production, and other important

Abbreviations: ADM, Actuator disk method; ALM, Actuator line method; ANN, Artificial neural network; CFD, Computational fluid dynamics; CNN, Convolutional neural networks; DNS, Direct numerical simulation; HAWT, Horizontal-axis wind turbine; LES, Large eddy simulation; ML, Machine learning; NREL, National Renewable Energy Laboratory; PALM, Parallelized Large-Eddy Simulation Model; RANS, Reynolds averaged Navier Stokes; VAWT, Vertical-axis wind turbine; XGBoost, eXtreme Gradient Boosting.

* Corresponding author.
E-mail address: mahdi.nakhchi@northumbria.ac.uk (M.E. Nakhchi).

https://doi.org/10.1016/j.enconman.2023.117708
Received 17 May 2023; Received in revised form 20 September 2023; Accepted 24 September 2023
Available online 30 September 2023

Table 1
Recent studies on power and wake prediction of wind farms using machine learning.

Author	ML method	Main findings
Zhang and Zhao [16] (2020)	Deep learning	• Wind farm wake prediction. 4.8 % error compared to analytical models.
Brogna et al. [27] (2020)	Gradient-based and Gradient-free methods	• Optimization of wind farms in complex terrain under uncontrolled conditions
Ti et al.[17] (2020)	ANN model based on the backpropagation (BP) algorithm	• Wind farm wake prediction without yaw misalignment. Computation time could grow exponentially for wind farms.
Ti et al. [28] (2021)	ANN	• Power and wake profiles of wind farms are predicted with uncontrolled wake conditions. Computational expense is high for large-scale wind farms. RANS method is used for collecting wake profiles.
Chen et al. [29] (2021)	Multilayer perceptron neural network (MLP-NN)	• Wake profile and power production of multiple turbines were predicted using low-fidelity RANS model.
Yang et al. [19] (2022)	ANN	• power estimate with a mean error of 1.13 %. Requires significant computation recourses for training and testing.
Purohit et al. [24] (2022)	SVR, ANN, and XGBoost	• Wake profile of a single NREL-5 MW wind turbine predicted. XGBoost requires less training time and more accurate. Yaw-misalignment was not considered.

parameters. The Parallelized Large-Eddy Simulation Model (PALM) [10] is also used to compare the ADM and ALM approaches in wind farm models. However, this model is currently computationally expensive

Table 2
Geometric parameters of the horizontal-axis wind turbine.

Parameter	Value
Rated output power (P_{out})	5 MW
Rated wind speed (U_{rated})	11.4 m/s
Tip-speed ratio	7.0
Rotor diameter (D)	126.0 m
Rotational speed (Ω)	12.1 rpm
No. of blades	3
Hub height	90.0 m
Hub diameter	3.0 m
Cone angle (β)	2.5°
Yaw angle (θ)	−30° to 30°

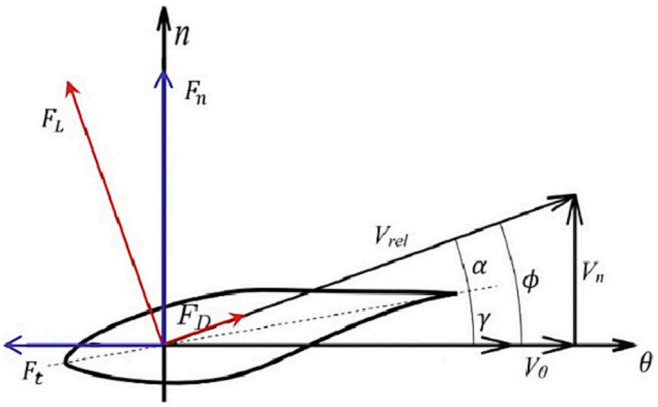
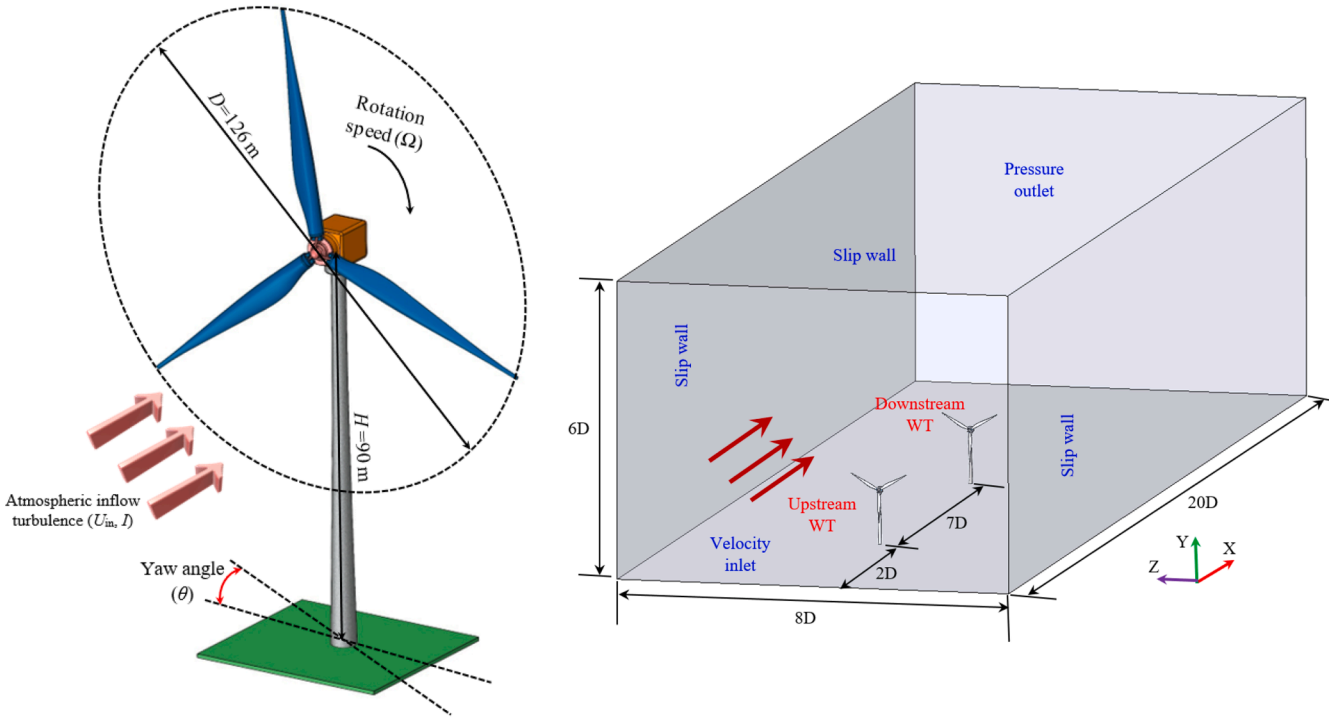


Fig. 2. Forces on the wind turbine mid-section using ALM method.



a) Wind turbine model b) Computational domain

Fig. 1. Schematic view of NREL 5 MW wind turbine under yaw-control conditions.

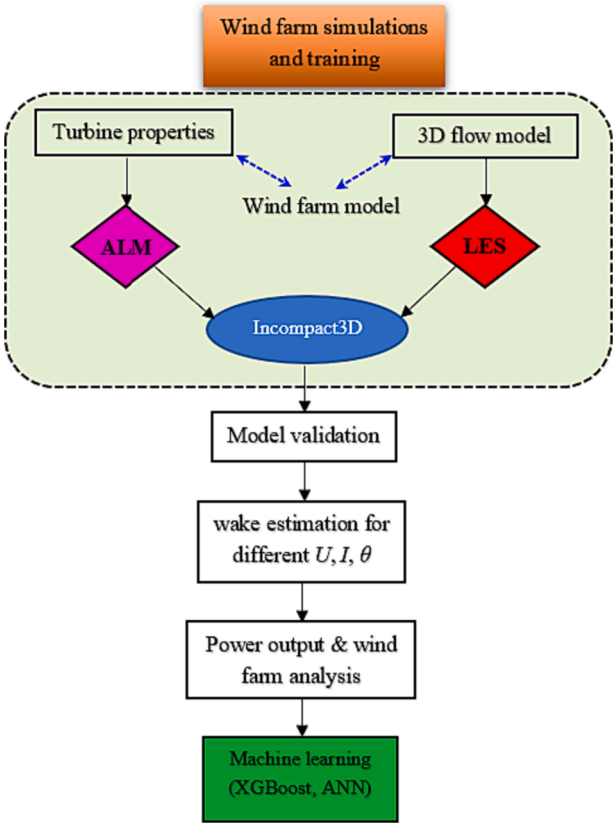


Fig. 3. Wind farm simulation flowchart using ALM-LES method.

and too complex in the performance prediction of large wind farm layouts by considering the realistic inflow conditions under various wake conditions.

In the past few years, machine learning methods like artificial neural networks (ANN) [11] and convolutional neural networks (CNN) [12,13] have been used to predict mechanical problems in many different fields. Wind turbine researchers are now focusing more and more on machine learning techniques for both steady [14] and transient fluid flow [15,16]. Neural networks may be trained to predict the output parameters as a function of input parameters precisely. Ti et al. [17] used ANN architecture with 2000 sub-models with one hidden layer to predict the performance of a wind farm layout with different numbers of horizontal-axis wind turbines (HAWTs). For the training, they used the back-propagation (BP) algorithm. Biswas et al. [18] used ANN for the prediction of the wake region of vertical axis wind turbines (VAWTs) to examine the overall performance, power production and torque forces on the blade structures [18]. Yang et al. [19] combined ANN with a Bayesian machine learning algorithm to predict the wake of yaw-controlled wind farms. To directly estimate the power production of a wind farm by considering the wake losses, machine learning methods like the Long Short Term Memory (LSTM) network [20] are used. The review study of Stetco et al. [21] about the applications of machine learning methods to predict the performance of wind turbines showed that it is challenging to estimate the turbine wake, which needs to know the link among inflow circumstances, rotor properties, turbulent flow structure and vortex generations in the wake area. They showed that accurate wake predictions of wind farms could be done much faster and with accuracy comparable to high-fidelity CFD simulations if an appropriate machine learning model were used. However, the previous machine learning studies on wake prediction of wind farms require significant amount of time and also many simplifications must be considered to the wind farm model to make it possible to apply machine learning with significant amount of data.

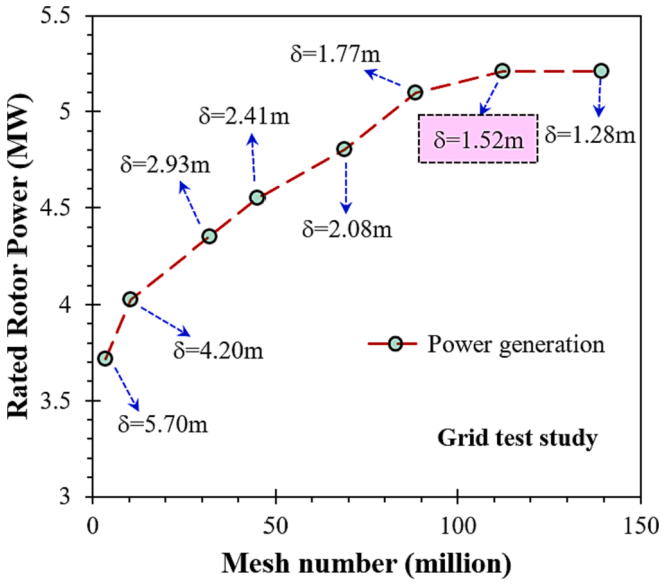


Fig. 5. Grid independence study over single wind turbine (U_{in} = 11.4 m/s, I = θ = 0).

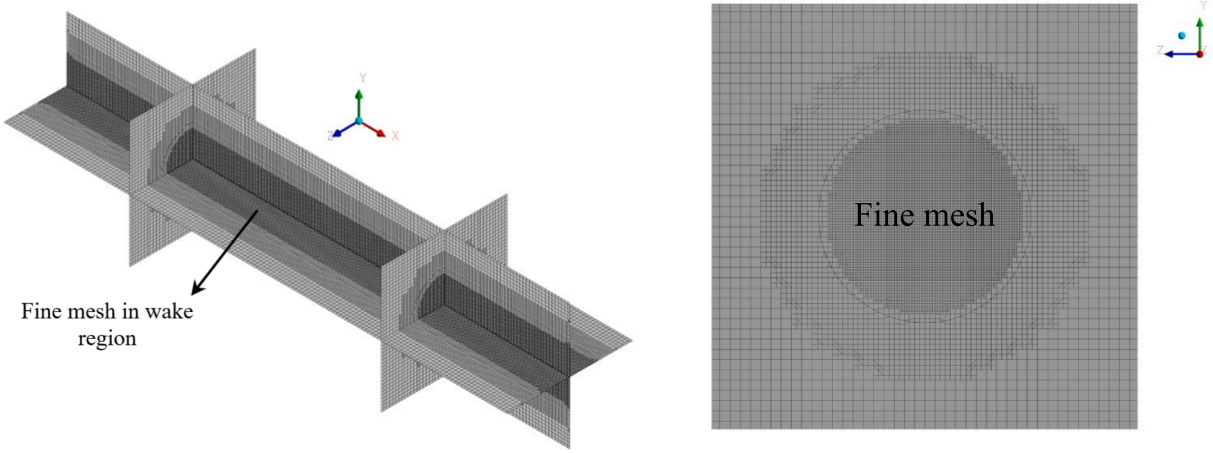


Fig. 4. Mesh generation over two wind turbines in tandem configuration.

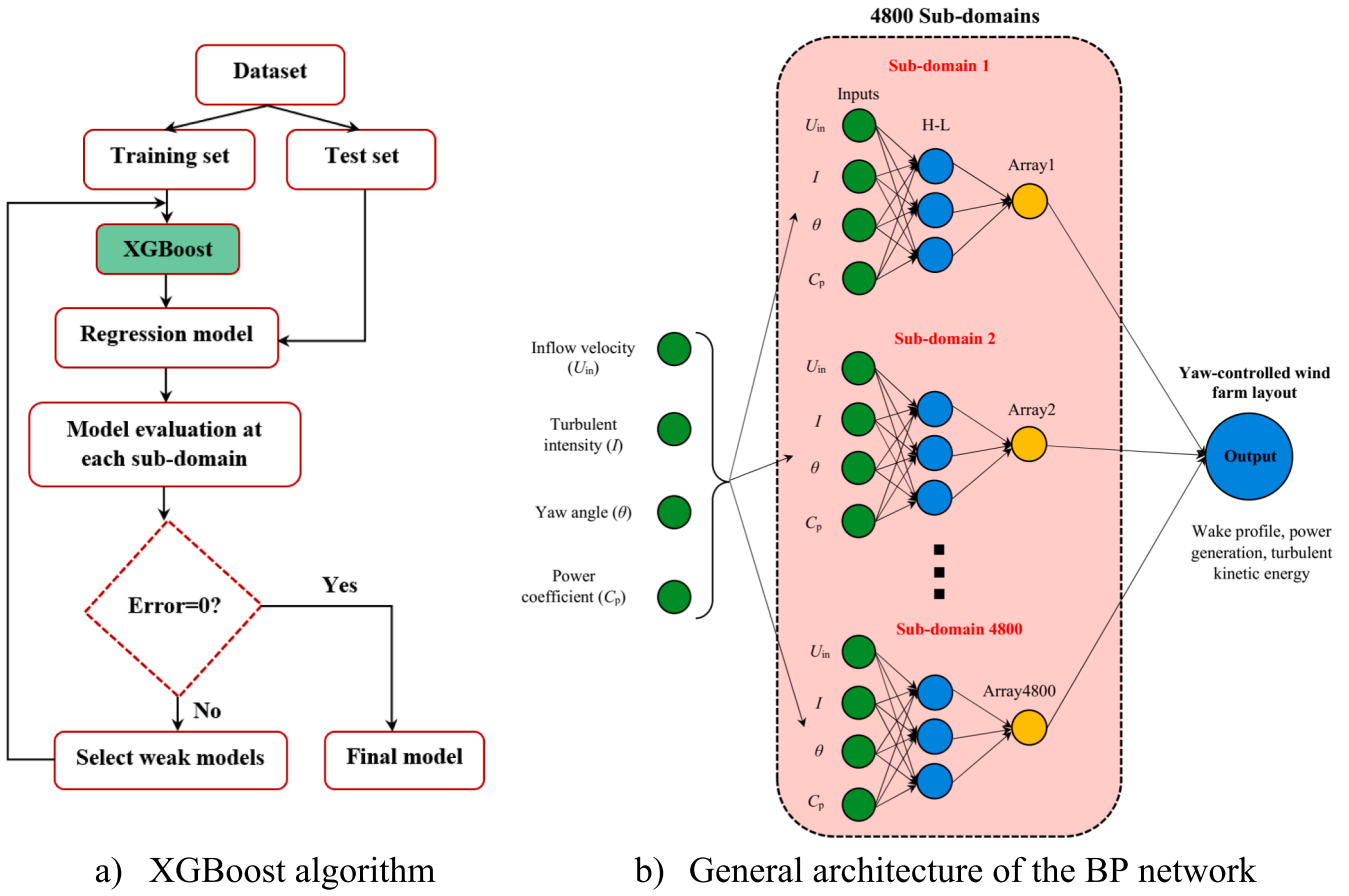


Fig. 6. General architecture of XGBoost algorithm and BP network for machine learning over wind farm layout.

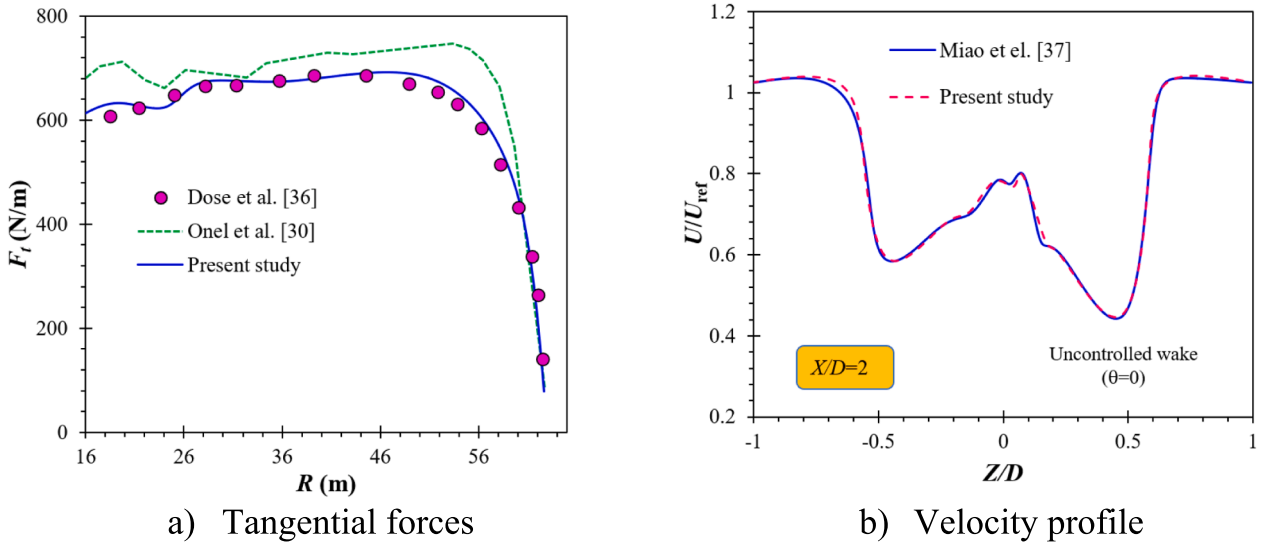


Fig. 7. Validation of the present model with previous models over NREL 5 MW wind turbine.

The eXtreme Gradient Boosting (XGBoost), is a modified and efficient machine learning algorithm which is an improved edition of the Gradient Boosting framework [22,23]. The main improvement is in the regularised loss functions in the objectives of XGBoost, which improves its efficiency and performance for training large amounts of data. XGBoost is an effective collective machine learning algorithm which merges the output of particular output trees into a joint output. A recent study by Purohit et al. [24] showed that XGBoost is significantly more efficient compared to the artificial neural network (ANN) and support vector regression (SVR) machine learning algorithms in velocity and turbulence intensity prediction of wind turbine wakes. In addition, XGBoost has been successfully used to detect the faults in wind turbines [25] and to predict the power production of wind power [26]. The summary of the previous machine learning studies on power and wake prediction of different wind farm layouts is provided in Table 1.

Based on the above literature review, wake prediction has been the

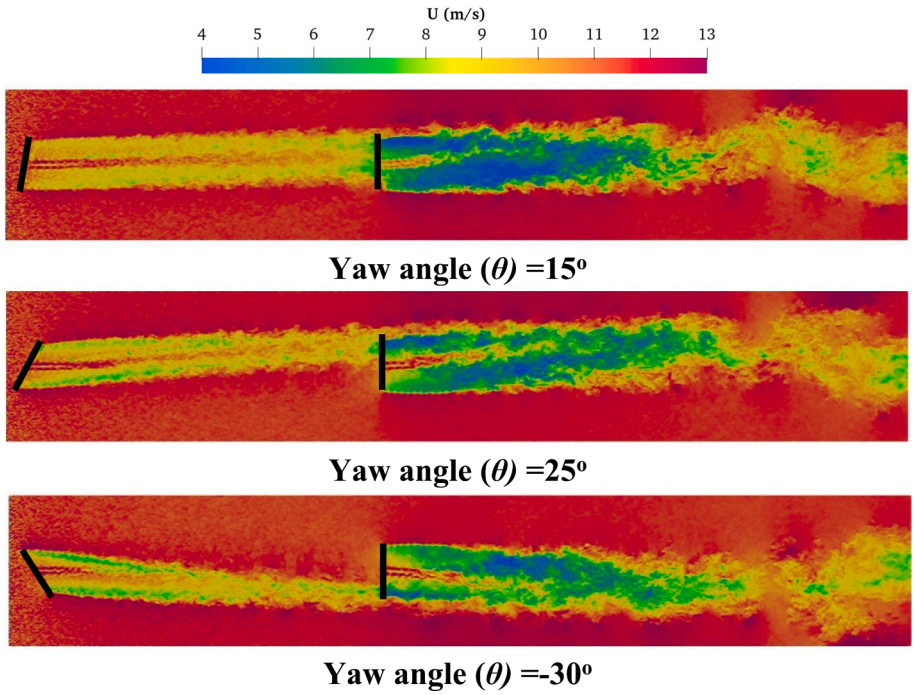


Fig. 8. Instantaneous axial velocity contours for different yaw angles.

subject of a relatively small amount of research through the application of machine learning approaches. Moreover, all of the previous studies were focused on typical straight wind turbines in their training models. However, in real wind farms, the upstream wind turbines are in wake-controlled conditions to increase their overall power production. The main reason that the previous machine learning studies on wind farms have imposed several simplifying assumptions was that the training time and precision would be significantly affected by considering more realistic input parameters. Recently, Ti et al. [17] proposed a machine learning analysis to model the wake profile of a wind farm layout in tandem form. They used ANN based on back-propagation (BP) for this purpose. However, they used a low-fidelity model for CFD simulations (ADM-RANS) to reduce the computational costs. In addition, they used wind turbines in uncontrolled wake conditions. Consequently, it is crucial to suggest a novel machine learning investigation that relies on advanced techniques such as LES and takes into account the influence of wake misalignment, such as yaw control, within the training algorithm. Reducing the training time is also another major parameter in machine learning based on a huge amount of data. In summary, the main novelties of this work can be summarised as:

- Considering the yaw angle misalignment effects in machine learning of a wind farm at different inflow velocities for the first time.
- Using XGBoost to predict wind farm wake and power generation for the first time.
- Significant reduction in error and training time compared to ANN and deep learning algorithms proposed by previous studies.
- Saving tens of thousands of CPU hours on a high-performance computing cluster.

The paper's structure highlights the key sections: 1. Introduction, literature review on the significance of machine learning in wind farms. 2. Physical description of the wind farm layout. 3. Mathematical formulation of the LES method using the Incompact3D code. Additionally, we will introduce the optimisation equations of the XGBoost method. 4. Results and discussion, where we analyse the ML outcomes and compare the precision of the proposed model, and 5. Conclusion, summarising the findings and emphasising their broader implications.

2. Physical description

The computation domain of two NREL 5 MW wind turbines by considering the inflow disturbance is shown in Fig. 1. The wind turbine characteristics are also provided. The wind turbines have a rotor diameter of D = 126 m and rotate with a rotational speed of Ω = 12.1 rpm. The domain has a 20D length in the streamwise direction (x), 8D in the lateral direction (z) and a height of 6D in the y-direction. The upstream wind turbine is placed at a 2D distance from the inlet, and the distance between the wind turbines is kept constant at 7D [30]. The typical distance (7D) between the wind turbines is used to avoid additional wake at smaller distances and additional power transmission and maintenance costs of larger wind farms with more distance (larger than 7D) between the turbines. This distance also provides the possibility of validation with previous studies in this field. The inflow turbulent intensity varied from 2 % to 26 % in this study, and the inflow velocity in the streamwise direction (U) varies from 11.4 to 15 m/s. The contours and simulation results are provided for the rated inflow velocity of 11.4 m/s over the NREL 5 MW wind turbine to produce the maximum rated power. The physical parameters of the 3D wind farm used in this study are provided in Table 2. The yaw angle of the upstream turbines varies between $(-30^o < \theta < 30^o)$. It is mentioned by Nash et al. [1] that the yaw angle could be positive or negative. The velocity inlet boundary condition is selected at the inlet, while the pressure outlet boundary condition is selected at the outlet of the domain. Slip wall boundary condition is imposed on the sidewalls of the domain. It should be noted that using higher yaw angles will impose additional forces on the wind turbine blades and will reduce their life expectancy. However, in this study all possible yaw angles are considered to provide a general machine learning approach for various wake control conditions.

3. Mathematical formulation

The Incompact3D code is a high-order finite-difference fluid flow solver on a Cartesian mesh. This code is developed in Fortran language to model different incompressible flows [31]. The main advantage of this solver is its better simulation speed and convergency with millions of elements on HPC clusters, and it provides the ability to model wind

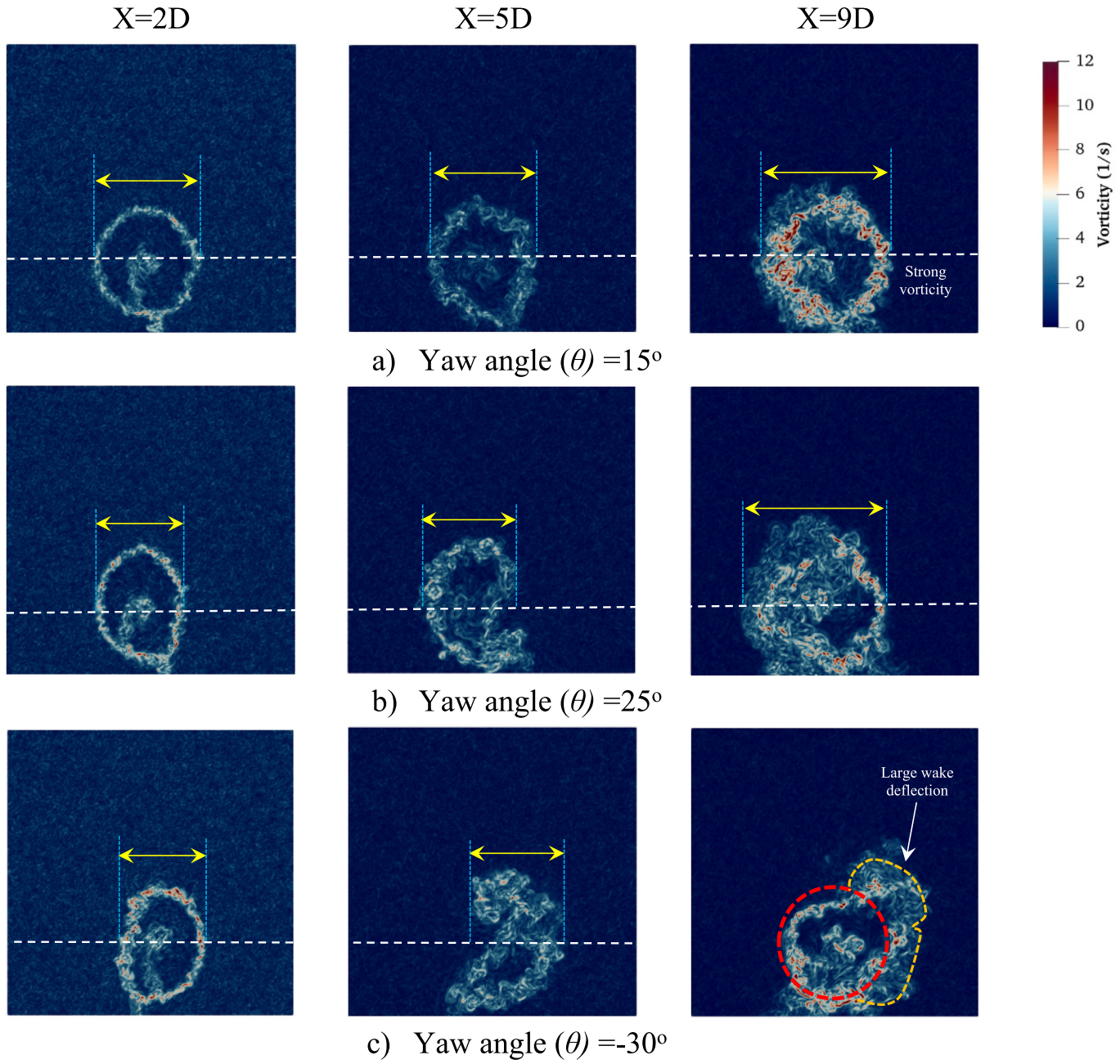


Fig. 9. Cross-section view of instantaneous vorticity magnitude in the wake region of the wind farm layout at different sections.

farms by considering all the geometrical parameters of the wind turbines, such as all details of wind turbine airfoils, tower sections, hub height, and also yaw, tilt and cone angles. Moreover, it can model both horizontal-axis and vertical-axis wind turbines with a different number of blades. This code can predict essential output parameters of wind farms, such as the overall output power.

The solver provides the ability to select direct numerical simulation (DNS) or LES for the transient flow simulation in the domain. In this study, the LES method with a time step size of $\delta t = 0.02sec$ is chosen for the simulations. This method can be used as a high-resolution turbulence model for the present simulations to capture both instantaneous and time-averaged results in the wake region of the wind turbines. To capture small eddies, a subgrid-scale (SGS) model is selected for the LES method. The governing equations of this model can be expressed as [32]:

$$\frac{\partial \widetilde{u}_i}{\partial x_i} = 0 \tag{1}$$

$$\frac{\partial \widetilde{u}_i}{\partial t} + \widetilde{u}_j \frac{\partial \widetilde{u}_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial \widetilde{p}}{\partial x_i} + \nu \frac{\partial}{\partial x_j}\left(\frac{\partial \widetilde{u}_i}{\partial x_j}\right) - \frac{\partial \tau_{ij}^{SGS}}{\partial x_j} + s_i \tag{2}$$

Where u, p, and ν are velocity tensor, pressure, and kinematic viscosity, respectively. s_i is a source term, and $\tau_{ij}^{SGS} = \widetilde{u_i u_j} - \widetilde{u}_i \widetilde{u}_j$ is the SGS symmetric stress tensor. In the present work, the Smagorinsky model is used for the numerical modelling of the wind farm domain. Based on the Smagorinsky model, the subgrid-scale stress tensor can be expressed as [32]:

$$\tau_{ij}^{SGS.dev} = \tau_{ij}^{SGS} - \frac{1}{3}\tau_{kk}^{SGS}\delta_{ij} \approx -2\nu_{SGS}\widetilde{S}_{ij} \tag{3}$$

where $\widetilde{S}_{ij} = \frac{1}{2}\left(\frac{\partial \widetilde{u}_i}{\partial x_j} + \frac{\partial \widetilde{u}_j}{\partial x_i}\right)$ is the resolved strain tensor. The SGS viscosity can be written as [32]:

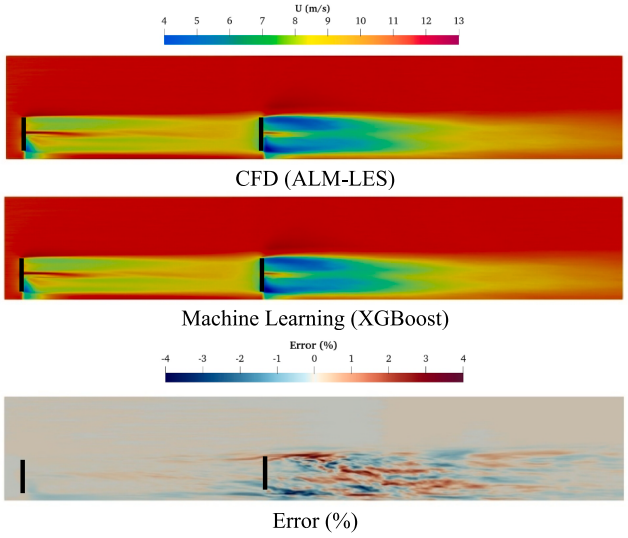


Fig. 10. Front view of comparison between XGBoost and ALM-LES results of axial wake generation of two wind turbines under yaw-controlled conditions ($\theta = 15^{\circ}$).

$$\nu_{SGS} = (C_s\Delta)^2\left|\widetilde{S}_{ij}\right| \tag{4}$$

In which $C_s = 0.168$ is the Smagorinsky's constant. The resolved strain magnitude can be evaluated as $\left|\widetilde{S}_{ij}\right| = \sqrt{2\widetilde{S}_{ij}\widetilde{S}_{ij}}$.

To model the wind turbine blades, hub and tower, the actuator line method [33] is used to replace each part of the wind turbine with body force lines and link them together with some actuator points. The details of actuator points employed on the NREL wind turbine blades can be found in [34]. Fig. 2 shows the details of forces and different design parameters at the mid-section of the NREL 5 MW turbine. V_n and V_θ are the normal and tangential velocity components on the aerofoil, γ is the pitch angle, $\phi = \tan^{-1}\left(\frac{V_n}{r\omega - V_\theta}\right)$ is the velocity angle, and $\alpha = \phi - \gamma$ is the angle of attack. The relative and tangential velocity components (V_{rel}, V_t) are calculated as [35]:

$$V_{rel} = \sqrt{V_n^2 + V_t^2},\ V_t = r\omega - V_\theta \tag{5}$$

The drag and lift forces (F_D,F_L) on wind turbine airfoils can be computed as [35]:

$$F_L = \frac{1}{2}\rho V_{rel}.c.C_L \tag{6}$$

$$F_D = \frac{1}{2}\rho V_{rel}.c.C_D \tag{7}$$

where c is the chord length of the blade in streamwise direction ρ, C_L and C_D are density, drag and lift coefficient, respectively. To calculate the forces on the wind turbine blades in the normal and tangential directions, the lift and drag forces can be divided as:

$$F_n = F_L cos\phi + F_D sin\phi,\ F_t = F_L sin\phi - F_D \cos\phi \tag{8}$$

The drag and lift forces on each cell of the turbine tower can be calculated as

$$F_D^{tower} = 0.5\rho U_{ref}^2 C_{D_{tower}} A_{cell} \tag{9}$$

where $C_{D_{tower}} = 1.0$ is chosen as the tower induced drag coefficient.

The CFD simulation flowchart of the ALM-LES method used in this study using the Incompact3D solver is provided in Fig. 3. For different inflow velocities and turbulent intensities, the simulations were

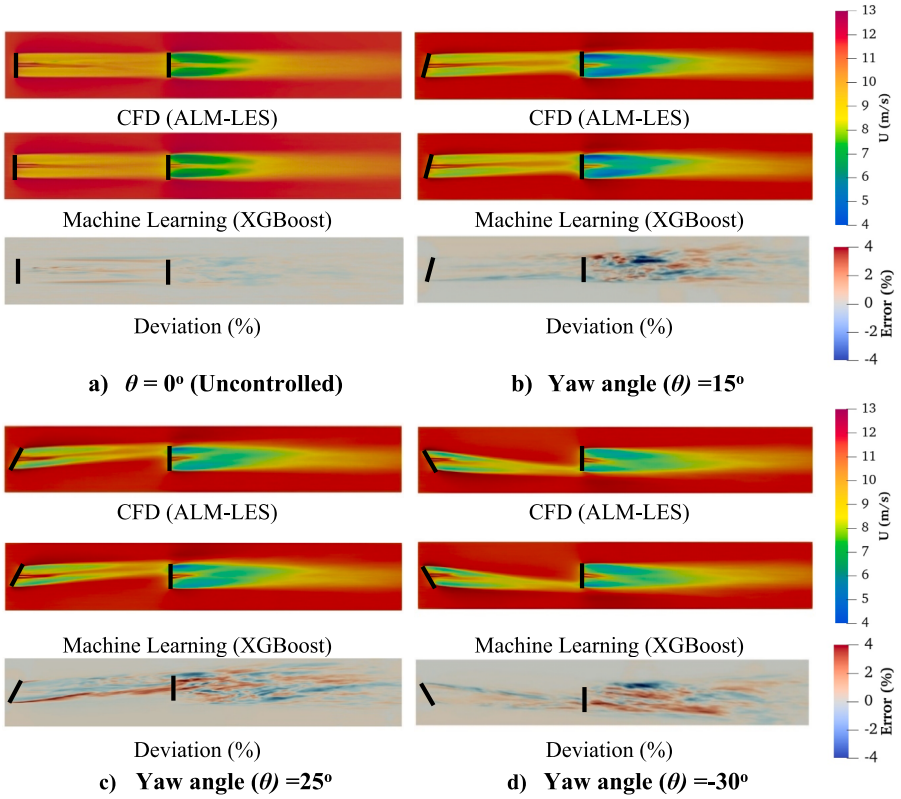


Fig. 11. Hub-height view of comparison between XGBoost and ALM-LES results of axial wake generation of two wind turbines under uncontrolled and yaw-controlled conditions.

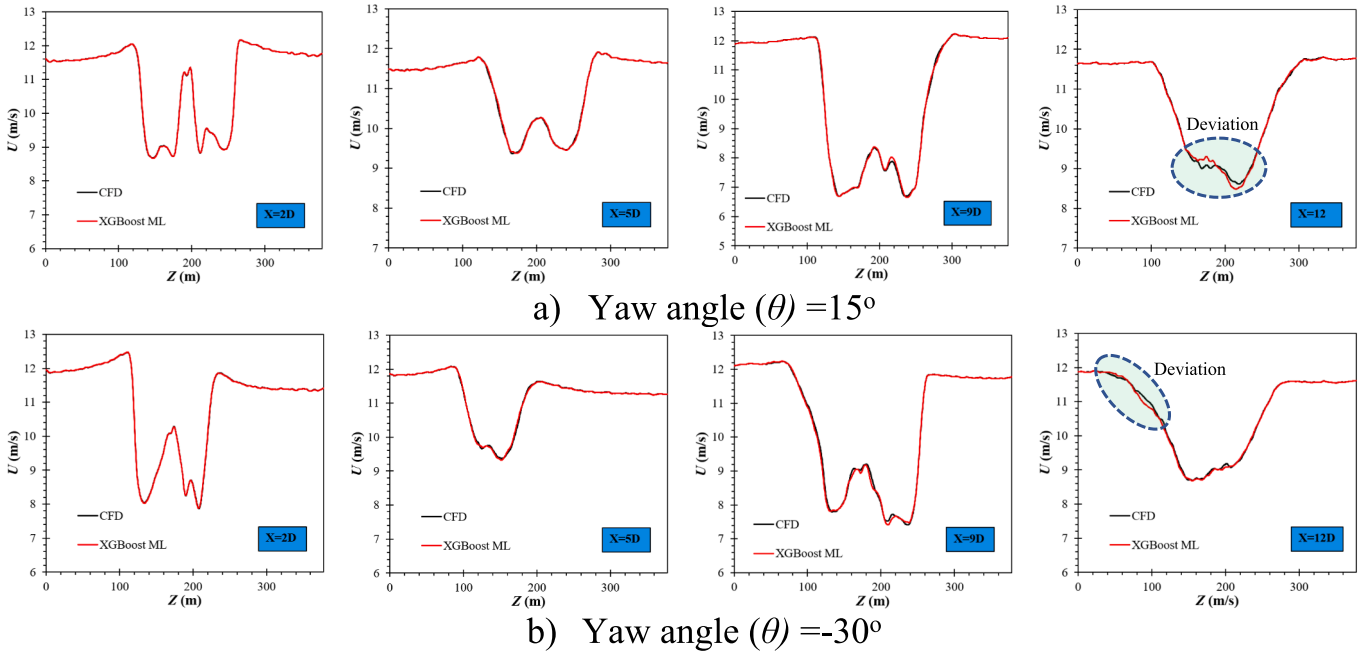


Fig. 12. Comparison between CFD and ML for wake profiles at different distances from downstream yaw-controlled wind turbine.

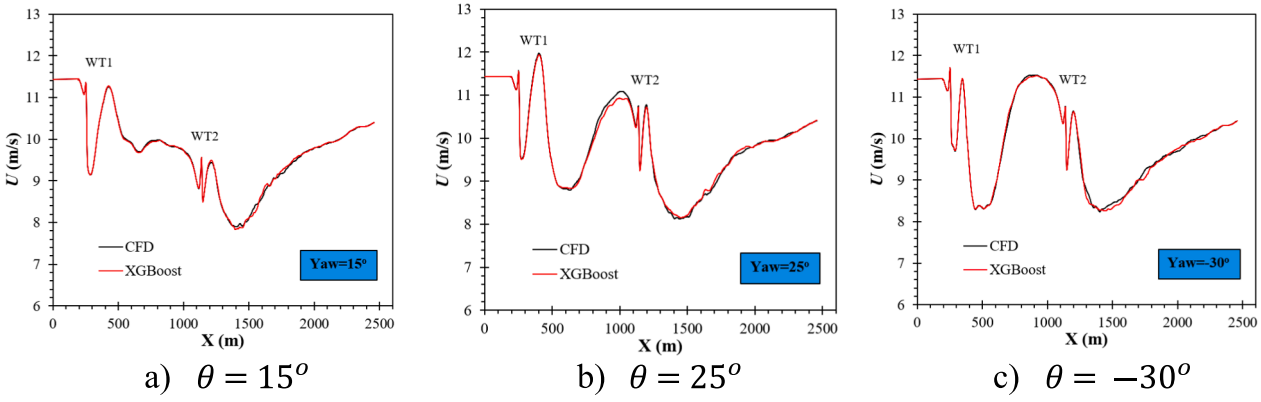


Fig. 13. Comparison of the CFD results with machine learning analysis for axial velocity at the hub-height (Y = 90 m).

performed at different yaw-controlled conditions to capture the instantaneous and time-averaged wake profile in the downstream region of a wind turbine. The turbine properties, including the blade airfoils, hub, and tower, are imposed in the ALM section of the solver. The flow characteristics and boundary conditions are set in the LES part of the solver to model the turbulent flow structure in the domain. These data are linked together in the solver. After collecting all required data for different inflow velocities and yaw angles, the machine learning analysis is performed to predict the power generation and wake profile of the wind farm.

Cartesian mesh is chosen as the computational domain to perform the ALM-LES simulations. The mesh is refined at the centre to capture the wake and vorticities in the downstream region accurately. The mesh is gradually smoothed toward the other layers of the domain. Details of the mesh used in the wake region in this study are shown in Fig. 4.

A grid independence study is performed to find the optimum grid for the simulations. As shown in Fig. 5, eight different grids ($1.28 \leq \delta \leq 5.70m$) are chosen in this study. The grids illustrate the mesh size at the centre of the refined mesh region. The grid study shows that by raising the number of elements from 3.42 million to 10.36 million, power generation will be increased significantly. By gradually increasing the number of elements, this deviation becomes smaller. It can be seen that the deviation of output power for $\delta = 1.28m$ compared to $\delta = 1.52m$ (element number of 112.37 million) is less than 0.2 %. Therefore, this mesh size is accurate enough to capture the power production of the NREL 5 MW wind turbine.

In this study, the XGBoost algorithm is utilised to create a fast and efficient model for the prediction of wake and power production of the wind farm model at different yaw-controlled conditions. The XGBoost algorithm employs the boosting ensemble learning basis to reduce errors and create more accurate models [36]. Fig. 6a shows the XGBoost algorithm flowchart. The main advantage of XGBoost compared to the other machine learning methods is better accuracy in wind farm predictions and more efficient training in parallel computing. The training is performed on the UK national supercomputer (ARCHER2) on Intel CPU nodes. The input parameters are inflow hub-height velocity (U_{in}), turbulent intensity (I), yaw angle (θ), and thrust coefficient (C_t). The velocity data in the flow field are interpolated into 360,000 data points in a matrix **V** for the training process. The inflow velocity varies from 8 to 15 m/s. Turbulent intensity at the inlet of the wind farm varies from 2 to 26 % which is reported by previous studies over wind farm layouts [9,37]. The yaw angle, which is the main novelty of this study varies

Table 3
Comparison between the ANN and XGBoost machine learning results with CFD simulations for different yaw angles.

Yaw = 15°					
Axial location	CFD	XGBoost		ANN	
	U (m/s)	*U* (m/s)	Error %	*U* (m/s)	Error %
X = 500	10.442	10.415	−0.26	10.608	1.59
X = 1000	9.730	9.727	−0.03	9.280	−4.62
X = 1500	8.101	8.113	0.15	7.779	−3.97
X = 2000	9.800	9.799	−0.01	9.872	0.73
Yaw = 25°					
Axial location	CFD	XGBoost		ANN	
	U (m/s)	*U* (m/s)	Error %	*U* (m/s)	Error %
X = 500	9.381	9.385	0.04	9.572	2.03
X = 1000	11.083	10.924	−1.43	10.590	−4.44
X = 1500	8.179	8.217	0.46	8.014	−2.01
X = 2000	9.830	9.832	0.02	10.02	1.93
Yaw = -30°					
Axial location	CFD	XGBoost		ANN	
	U (m/s)	*U* (m/s)	Error %	*U* (m/s)	Error %
X = 500	8.327	8.334	0.08	8.327	−0.08
X = 1000	11.413	11.395	−0.15	11.108	−2.67
X = 1500	8.448	8.331	−1.38	8.187	−3.09
X = 2000	9.698	9.763	0.67	9.580	−1.22

from −30 to 30 degrees. The wind turbines usually operate in smaller yaw misalignments. However, large angles are selected for training purposes to make the predictions more general. In this study, 32 ALM-LES sample simulations are required to capture all necessary data for the training. The objective function at *t*-th iteration to be trained in XGBoost can be expressed as [24]:

$$\mathscr{L}^t = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{t-1} + f_t(x_i)\right) + \Omega(f_t) \tag{10}$$

where y_i is the target parameter at i th case, $\hat{y}_i^{t-1}$ is the prediction of the parameter at $t-1$ iteration, f_t is a function that minimises the objective function $\mathscr{L}^t$, and $\Omega(f_t)$ is the regularisation term. By applying a second-order Taylor expansion in the XGBoost algorithm, this equation can be expressed as [24]:

$$\widehat{\mathscr{L}}^t = \sum_{i=1}^{n} \left[g_i f_t(x_i) + 0.5 h_i f_t^2(x_i)\right] + \Omega(f_t) \tag{11}$$

where g_i and h_i are the gradients and Hessians of the cost function, respectively. The following learner minimises the loss function by calculating a scoring function (Eq. (12)) which evaluates the training tree quality, and the optimal weights can be written as:

$$\widehat{\mathscr{L}}^t(q) = -0.5 \sum_{j=1}^{T} \frac{\left(\sum i \in I_j g_i\right)^2}{\sum i \in I_j h_i + \lambda} + \gamma T \tag{12}$$

$$w_j^* = \frac{\sum i \in I_j g_i}{\sum i \in I_j h_i + \lambda} \tag{13}$$

The exact stingy algorithm was utilised to compute all possible trees for all parameters to find the optimum tree structure. The loss reduction after splitting can be expressed as:

$$\mathscr{L}_{splt} = 0.5 \left[\frac{\left(\sum i \in I_L g_i\right)^2}{\sum i \in I_L h_i + \lambda} + \frac{\left(\sum i \in I_R g_i\right)^2}{\sum i \in I_R h_i + \lambda} + \frac{\left(\sum i \in I g_i\right)^2}{\sum i \in I h_i + \lambda} \right] \tag{14}$$

The machine learning architecture of ANN based on the BP algorithm which is used for comparison with XGBoost in this study is shown in Fig. 6b. More details about this algorithm and its application for wake prediction of wind farms can be found in [17].

4. Results and discussion

The numerical method used in the present analysis is validated before further analyses to ensure that the results are accurate and reliable. As there are no physical experiments for this analysis, the results are first validated with various numerical studies. The variation of the tangential forces with the blade radius is computed from the present simulation and compared to those of Dose et al. [38] and Onel et al. [32], as shown in Fig. 7a. It is observed that a close agreement is obtained between the present study and the fully resolved mesh (FRM) analysis of Dose et al. [38], whereas there are some deviations among the present simulation and ALM results of Onel et al. [32]. This is due to different numerical methods employed; however, the deviations are very small and within an acceptable level. It should also be noted that the present simulation captures the tip loss behaviour reasonably well, which is very important for wind turbine analysis. In addition to the forces applied to the blades, it is also important to capture the wakes. Fig. 7b compares the dimensionless wake profile, expressed in U/U_{ref}, extracted at 2D (D is the rotor diameter) from the turbine in the downstream wake region, computed from the present simulation and that of Miao et al. [39]. The two studies show a remarkable level of agreement, which ensures that the numerical model can capture the downstream wake profiles.

The effect of the yaw angle misalignment on the unsteady flow behaviour around two wind turbines in arrays is illustrated in Fig. 8

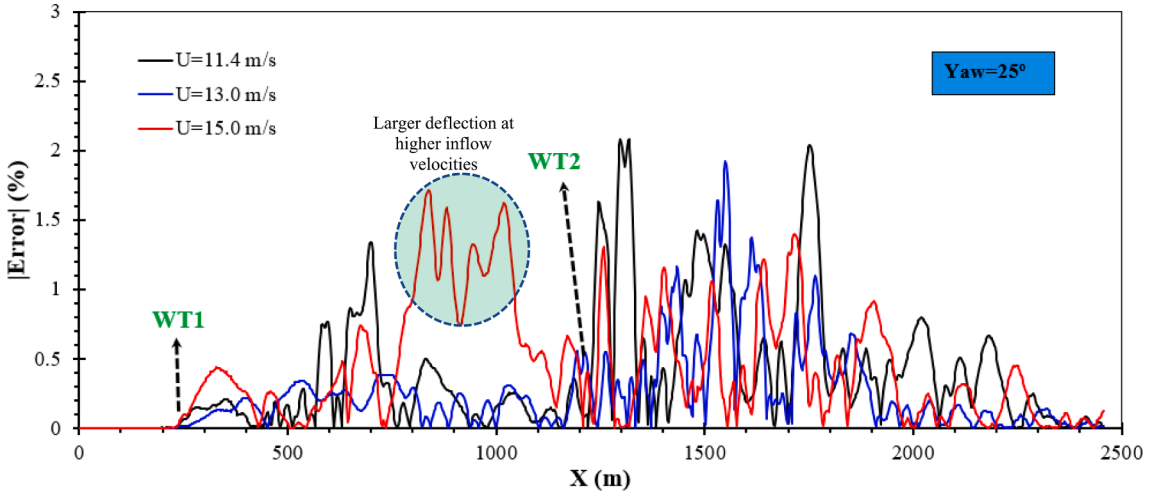


Fig. 14. The effects of Inflow velocity on prediction accuracy of machine learning over yaw-controlled wind turbines.

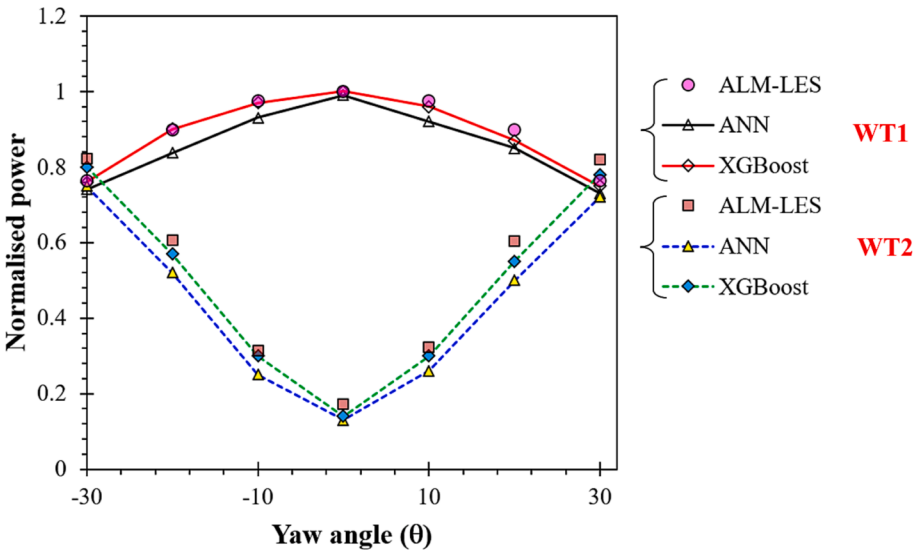
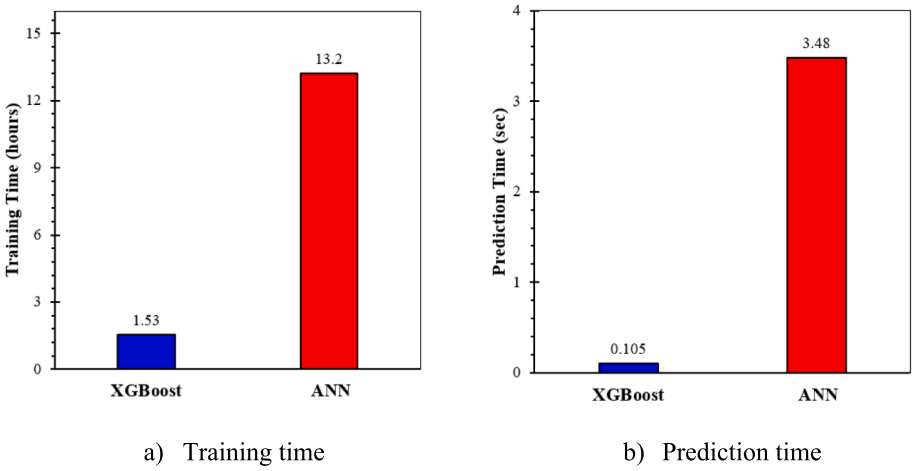


Fig. 15. Comparison of normalised power generation of two wind turbines for ML with CFD at different yaw angles.



a) Training time b) Prediction time

Fig. 16. Comparison of training and prediction time of ANN and XGBoost over wind farm layout.

through instantaneous velocity contours. It can be clearly seen that the wake close to the second turbine is highly influenced by the upstream turbine at all yaw angles shown in this figure. At $\theta = 15^{\circ}$, the wake of the upstream turbine is slightly shifted to the side due to the yaw misalignment to the inflow wind before reaching the downstream turbine. Despite the wake deflection, the downstream wind turbine is still majorly in the wake of the upstream turbine. As a result, the velocity magnitudes are much lower and the vortex generations are stronger in the wake region of the downstream turbine compared to that of the upstream turbine. These effects on the downstream wind turbine can be reduced by controlling the yaw angle. Therefore, the performance of the downstream wind turbine can be improved. The wakes behind the downstream turbine at these angles are also strong due to the influences of the turbulence from the upstream turbine; however, the velocity fields are slightly improved compared to the $\theta = 15^{\circ}$ case.

The evolution of vorticity behind each turbine at different distances (X = 2D, 5D and 9D) are plotted on a plane normal to the wind direction in Fig. 9. At X = 2D (2D behind the upstream wind turbine), a tip vorticity ring can still be observed at the yaw angle of 15°. However, this vorticity is deformed to the side when the yaw angle is further increased. The side to which the wake is deflected depends on whether the yaw angle is controlled in the positive direction (clockwise looking at from the wind direction) or negative direction (anti-clockwise looking at from the wind direction). The wake deflections are more pronounced behind the downstream wind turbine at X = 5D and X = 9D as the vortex generations are stronger as seen in the previous figure. At $\theta = 25^{\circ}$ and $\theta = -30^{\circ}$, the wakes of the downstream wind turbines at both distances are completely deformed. The vorticity close to the downstream turbine is stronger than that of the upstream turbine, and the magnitude becomes higher as it goes further downstream behind the downstream wind turbine.

Fig. 10 compares the time-averaged velocity distributions across wind turbines obtained from a flow simulation using an ALM-LES method and predicted from the Machine Learning procedure using the XGBoost tool. An error contour is also presented to indicate the difference in percentage between the two methods. The wake development from each turbine and the wake interactions can be observed in this figure. The velocity magnitudes behind the downstream turbine are significantly lower compared to that of the upstream turbine. It is seen that the flow fields predicted by the two methods are very similar, which indicates that the XGBoost machine learning model can capture and predict the wake structures generated from both upstream and downstream turbines correctly. The errors between these two methods are within 4 %, which is an acceptable limit. It is also the first time that the effects of the yaw angle misalignment are considered in a machine learning model. Unlike the study of Ti et al. [17], the present study uses

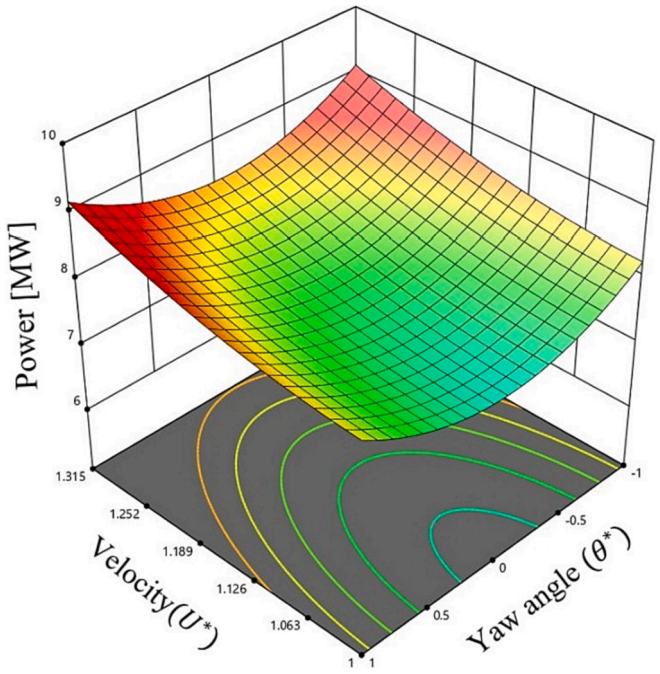


Fig. 17. Overall power estimation of the wind farm at different yaw angles and inflow velocities using the XGBoost algorithm.

an LES method to solve the flow governing equations to train the machine learning model. At the same time, it improves the resolution and accuracy of the unsteady flow structures predicted by the machine learning model.

The influences of the yaw angle variations on the velocity fields at the hub height and wake behaviour are demonstrated in Fig. 11. The yaw angles are varied between -30° and 30°. When the yaw angle is set to 0°, the downstream wind turbine operates entirely within the wake of the upstream turbine, leading to a considerable decrease in the performance of the downstream turbine. At $\theta = 15^{\circ}$, it can be observed that the wake is deflected to the left side, looking from the wind direction, which reduces the impact of the upstream wake on the downside one. Although the performance of the upstream turbine is affected by the misalignment of the yaw angle to the inflow wind, the power of the downstream wind turbine can be improved, thereby increasing overall power output. The upstream wake is further deflected as the yaw angle is increased. The comparisons between the ALM-LES simulations and the predictions of the XGBoost model show that the flow fields obtained from the two methods are almost identical at $\theta = 0^{\circ}$. Overall, the proposed machine learning model accurately predicted the unsteady flow behaviour within an acceptable error limit (±4%).

In addition to the velocity contours, the velocity profiles at various distances from the turbine in the downstream region are also plotted in Fig. 12 to quantitively compare the two methods. The results are plotted for the $\theta = 15^{\circ}$ and $\theta = -30^{\circ}$ cases. The profiles are almost identical in the wake region of the upstream one (X = 2D and X = 5D), but there are some deviations in the wake region of the downstream wind turbine (X = 9D and X = 12D), where the flow is more unsteady and turbulent due to the merging of wake structures. Nevertheless, the deviations are neglectable, and it is concluded that the machine learning model employed in this study well-predicted the wake profiles, which is essential for the aerodynamic analysis of wind turbines.

Fig. 13 illustrates the comparison between the CFD results with machine learning analysis for the axial velocity at the hub height of two wind turbines in tandem configuration. The comparison shows that the velocities at the inlet and wake of the first turbine are captured precisely for all cases. These deviations become larger at higher yaw misalignment angles. It observed that differences become slightly larger in the wake region of the second wind turbine, which is due to uncontrolled wake conditions of the downstream wind turbine.

Table 3 quantitatively compares the accuracy of XGBoost, the employed machine learning model, with the ANN machine learning model and the CFD simulation for the prediction of streamwise velocity magnitudes at different axial locations at various yaw angles. The errors higher than 1 % are observed at X = 1000 at $\theta = 25^{\circ}$ and X = 1500 at $\theta = 30^{\circ}$. However, the errors are within 2 % and can be considered low, especially compared to the ANN model of which errors are higher than 4 %. Therefore, it can be noted that the XGBoost model can predict similar streamwise magnitudes as the CFD model without requiring significant computational resources.

The effects of inflow velocity on deviations between the machine learning analysis and CFD simulations are provided in Fig. 14. The wind turbine under investigation is a 5 MW offshore wind turbine with a rated wind speed of 11.4 m/s. Generally, the wind condition at sea is more steady and less turbulent than the condition on land. Therefore, it is assumed that the wind turbine is operating at around the rated wind speed, and three inflow speeds of 11.4 m/s, 13 m/s, and 15 m/s are selected in this study. The comparison is provided for three different inflow velocities at the yaw angle of 25°. The error is defined as $|Error| = \left|\frac{U_{ML}-U_{CFD}}{U_{CFD}} \times 100\right|$. The comparison is provided at the flow stream direction at the hub height of the two HAWTs. The results show that generally, the XGBoost machine learning can precisely capture the velocity for inflow velocities at an extreme yaw angle of 25°. The same trend was observed by comparing the results for different yaw angles. It is observed that small deviations can be detected in the wake region of the first wind turbine. These deviations become larger and reach 2 % at higher inflow velocities. For all cases, these deviations become larger in the wake region of the second wind turbine. The main reason is that the second wind turbine is in uncontrolled conditions and the vorticities are larger. At far wake regions of the second wind turbine, the deviations gradually start to decrease and reach below 1 % at far distances of X > 2000 m.

Fig. 15 illustrates the comparison of the normalised power generation of each wind turbine at various yaw angles, obtained from the ALM-LES method and the XGBoost machine learning method. The results predicted from the artificial neural network (ANN) method are also added to the comparison to evaluate the accuracy of each method. It can be seen that the predictions from the XGBoost machine learning method are in close agreement with the ALM-LES method, whereas some small offsets are present between the ALM-LES method and the ANN method. Results obtained demonstrate that the power produced from the upstream wind turbine is maximum at $\theta = 0^{\circ}$; however, the power from the downstream turbine is minimal at this angle. Although increasing the yaw angle reduces the power output of the upstream turbine as it is not facing the inflow wind directly and loses a certain amount of kinetic energy due to the misalignment of the yaw angle, it improves the performance of the downstream turbine, which increases the overall power output of both wind turbines combined. It is found that the greatest production is obtained at $\theta = \pm 30^{\circ}$. The results illustrate that the XGBoost can predict the power with a mean deviation of 0.94 % for different yaw angles. These deviations between ANN predictions and CFD results are larger. ANN can estimate the power generation with a mean deviation of 2.15 % for various tested yaw angles.

As shown in Fig. 16, the proposed method can provide a set of high-resolution flow data at a significantly reduced computational cost. It was observed that the training time of the XGBoost was 87.5 % faster than ANN with the same amount of input data. The training process of 4800 sub-domains took around 13.20 h for ANN, while it took only 1.53 h for XGBoost. The prediction of the XGBoost is also much faster than the ANN architecture based on the BP algorithm. The wake data prediction of XGBoost is just 0.105 sec, while this time is 4.480 sec for the ANN model.

The overall power production of the wind farm at different yaw

angles and inlet velocities are provided in Fig. 17. The yaw angle and inflow velocities are plotted in dimensionless forms as $U^* = U/11.4$, and $\theta^* = \theta/30$. These dimensionless parameters are in the ranges of $-1 < \theta^* < 1$ and $1 < U^* < 1.315$. It can be seen that the power production reaches higher than 9 MW under yaw-controlled conditions with the highest inlet velocity of 15 m/s. The overall power generation is improved by 2.37 MW compared to the uncontrolled wake conditions with $U^* = 1$, and $\theta^* = 0$.

5. Conclusion

With advances in computing technology, machine learning algorithms are now widely used in various engineering applications. This paper utilises the benefits of the capabilities of a machine learning model to forecast the performance of two tandem wind turbines as well as to investigate the unsteady wake behaviour behind each turbine. A well-known ALM-LES method is employed to run the unsteady flow simulations to train the machine learning model. The XGBoost model is used as a machine learning model in this paper. Based on the results obtained, the following conclusions are drawn.

- The wake behind the upstream wind turbine is primarily deflected at the yaw angle of $\pm 30^{\circ}$, and the velocity field behind the downstream wind turbine is improved as the impact of the turbulence and wakes of the upstream wind turbine is significantly reduced due to wake deflection. At the axial distance of 9D behind the upstream wind turbine with the yaw angle of $\pm 30^{\circ}$, it can be clearly seen that the tip vorticity is completely deformed with stronger vortex generation.
- As a result, the overall power production from both the upstream and downstream wind turbines is improved as a combined power of 9 MW is achieved at $\theta = \pm 30^{\circ}$ and the inlet velocity of 15 m/s, which is 2.73 MW higher than the uncontrolled configuration.
- The errors between the ALM-LES method and the machine learning with the XGBoost method are within 4 %, which is acceptable. Errors higher than 2 % are associated with capturing the wakes of the downstream wind turbine with strong vortex generation. Nevertheless, a close agreement between the two methods is obtained in predicting the wake profiles, the axial velocity profiles and the flow field with small deviations detected at X = 9D and 12D.
- In terms of the training time, the XGBoost method is 87.5 % faster than the ANN method with the same amount of input data. The training process takes around 13.20 h with the ANN method, however; it takes only 1.53 h with the XGBoost method. The prediction of the wake data takes 0.105 sec with the XGBoost method, whereas it takes 4.480 sec with the ANN method. Despite faster training and prediction time, it is also observed that the XGBoost method can provide better accuracy compared to the ANN method.
- Results also illustrate that the XGBoost method predicts the power with a mean deviation of 0.94 % compared to 2.15 % with the ANN method.

One of the assumptions of this study is that the wind turbine under investigation operates at around the rated wind speeds, and only three wind speeds (11.4 m/s, 13 m/s, and 15 m/s) are considered in this analysis. In the future, a wide range of wind speeds below and above the rated wind speeds will be explored to investigate the effects of various wind speeds.

Future studies can focus on investigating the influence of atmospheric flows in machine learning of wind farm layouts. Also, the limitation of the machine learning study is the noticeable training time for large and more complex wind farm layouts. Therefore, it is essential to propose faster ML algorithms for this purpose. The authors will consider future works where blades yaw-angle and tilt-angle become part of the machine learning model. Including the HAWTs yaw-angle and tilt-angle in the machine learning model depends on the complexity of its implementation and collecting enough data for training and validation.

CRediT authorship contribution statement

M.E. Nakhchi: Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing – original draft. **S. Win Naung:** Conceptualization, Writing – review & editing, Investigation. **M. Rahmati:** Conceptualization, Writing – review & editing, Supervision.

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.

Data availability

Data will be made available on request.

References

[1] Nash R, Nouri R, Vasel-Be-Hagh A. Wind turbine wake control strategies: A review and concept proposal. Energy Conversion and Management 2021;245:114581.
[2] Li Z, Yang X. Large-eddy simulation on the similarity between wakes of wind turbines with different yaw angles. Journal of Fluid Mechanics 2021;921.
[3] Ge M, Zhang S, Meng H, Ma H. Study on interaction between the wind-turbine wake and the urban district model by large eddy simulation. Renewable Energy 2020;157:941–50.
[4] Su R, Gao Z, Chen Y, Zhang C, Wang J. Large-eddy simulation of the influence of hairpin vortex on pressure coefficient of an operating horizontal axis wind turbine. Energy Conversion and Management 2022;267:115864.
[5] Stevens RJ, Martínez-Tossas LA, Meneveau C. Comparison of wind farm large eddy simulations using actuator disk and actuator line models with wind tunnel experiments. Renewable Energy 2018;116:470–8.
[6] Draper M, Guggeri A, Mendina M, Usera G, Campagnolo F. A large eddy simulation-actuator line model framework to simulate a scaled wind energy facility and its application. Journal of Wind Engineering and Industrial Aerodynamics 2018;182:146–59.
[7] Qian G-W, Song Y-P, Ishihara T. A control-oriented large eddy simulation of wind turbine wake considering effects of Coriolis force and time-varying wind conditions. Energy 2022;239:121876.
[8] Van Der Laan MP, Sørensen NN, Réthoré PE, Mann J, Kelly MC, Troldborg N, et al. An improved k-∈ model applied to a wind turbine wake in atmospheric turbulence. Wind Energy 2015;18:889–907.
[9] Wu Y-T, Porté-Agel F. Modeling turbine wakes and power losses within a wind farm using LES: An application to the Horns Rev offshore wind farm. Renewable Energy 2015;75:945–55.
[10] Witha B, Steinfeld G, Heinemann D. High-resolution offshore wake simulations with the LES model PALM, Wind energy-impact of turbulence. Springer; 2014. p. 175–81.
[11] Marugán AP, Márquez FPG, Perez JMP, Ruiz-Hernández D. A survey of artificial neural network in wind energy systems. Applied Energy 2018;228:1822–36.
[12] Wang C, Xiong R, Tian J, Lu J, Zhang C. Rapid ultracapacitor life prediction with a convolutional neural network. Applied Energy 2022;305:117819.
[13] Shen S, Sadoughi M, Li M, Wang Z, Hu C. Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries. Applied Energy 2020;260:114296.
[14] Guo X, Li W, Iorio F. Convolutional neural networks for steady flow approximation. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining; 2016. p. 481–90.
[15] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 2019;378:686–707.
[16] Zhang J, Zhao X. A novel dynamic wind farm wake model based on deep learning. Applied Energy 2020;277:115552.
[17] Ti Z, Deng XW, Yang H. Wake modeling of wind turbines using machine learning. Applied Energy 2020;257:114025.
[18] Biswas A, Sarkar S, Gupta R. Application of artificial neural network for performance evaluation of vertical axis wind turbine rotor. International Journal of Ambient Energy 2016;37:209–18.
[19] Yang S, Deng X, Ti Z, Yan B, Yang Q. Cooperative yaw control of wind farm using a double-layer machine learning framework. Renewable Energy 2022;193:519–37.
[20] Göçmen T, Giebel G. Data-driven wake modelling for reduced uncertainties in short-term possible power estimation. Journal of Physics: Conference Series, IOP Publishing 2018;1037.
[21] Stetco A, Dinmohammadi F, Zhao X, Robu V, Flynn D, Barnes M, et al. Machine learning methods for wind turbine condition monitoring: A review. Renewable Energy 2019;133:620–35.

[22] Qiu R, Liu C, Cui N, Gao Y, Li L, Wu Z, et al. Generalized Extreme Gradient Boosting model for predicting daily global solar radiation for locations without historical data. Energy Conversion and Management 2022;258.
[23] Başakın EE, Ekmekcioğlu Ö, Özger M. Developing a novel approach for missing data imputation of solar radiation: A hybrid differential evolution algorithm based eXtreme gradient boosting model. Energy Conversion and Management 2023;280:116780.
[24] Purohit S, Ng E, Kabir IFSA. Evaluation of three potential machine learning algorithms for predicting the velocity and turbulence intensity of a wind turbine wake. Renewable Energy 2022;184:405–20.
[25] Zhang D, Qian L, Mao B, Huang C, Huang B, Si Y. A data-driven design for fault detection of wind turbines using random forests and XGboost. IEEE Access 2018;6:21020–31.
[26] Demolli H, Dokuz AS, Ecemis A, Gokcek M. Wind power forecasting based on daily wind speed data using machine learning algorithms. Energy Conversion and Management 2019;198:111823.
[27] Brogna R, Feng J, Sørensen JN, Shen WZ, Porté-Agel F. A new wake model and comparison of eight algorithms for layout optimization of wind farms in complex terrain. Applied Energy 2020;259:114189.
[28] Ti Z, Deng XW, Zhang M. Artificial Neural Networks based wake model for power prediction of wind farm. Renewable Energy 2021;172:618–31.
[29] Chen L, Wang H, Chin RJ, Luo H, Yao Y, Wu Z. An effective framework for wake predictions of tidal-current turbines. Ocean Engineering 2021;235:109403.
[30] Wang Y, Miao W, Ding Q, Li C, Xiang B. Numerical investigations on control strategies of wake deviation for large wind turbines in an offshore wind farm. Ocean Engineering 2019;173:794–801.
[31] Deskos G, Laizet S, Palacios R. WInc3D: A novel framework for turbulence-resolving simulations of wind farm wake interactions. Wind Energy 2020;23:779–94.
[32] Onel HC, Tuncer IH. Investigation of wind turbine wakes and wake recovery in a tandem configuration using actuator line model with LES. Computers and Fluids 2021;220:104872.
[33] Sorensen JN, Shen WZ. Numerical modeling of wind turbine wakes. Journal of Fluids Engineering 2002;124:393–9.
[34] Nakhchi M, Naung SW, Rahmati M. A novel hybrid control strategy of wind turbine wakes in tandem configuration to improve power production. Energy Conversion and Management 2022;260:115575.
[35] Ravensbergen M, Mohamed AB, Korobenko A. The actuator line method for wind turbine modelling applied in a variational multiscale framework. Computers and Fluids 2020;201:104465.
[36] Kiangala SK, Wang Z. An effective adaptive customization framework for small manufacturing plants using extreme gradient boosting-XGBoost and random forest ensemble learning algorithms in an Industry 4.0 environment. Machine Learning with Applications 2021;4:100024.
[37] Argyle P, Watson S, Montavon C, Jones I, Smith M. Modelling turbulence intensity within a large offshore wind farm. Wind Energy 2018;21:1329–43.
[38] Dose B, Rahimi H, Herráez I, Stoevesandt B, Peinke J. Fluid-structure coupled computations of the NREL 5 MW wind turbine by means of CFD. Renewable Energy 2018;129:591–605.
[39] Miao W, Li C, Pavesi G, Yang J, Xie X. Investigation of wake characteristics of a yawed HAWT and its impacts on the inline downstream wind turbine using unsteady CFD. Journal of Wind Engineering and Industrial Aerodynamics 2017;168:60–71.