Physics-Informed Neural Network surrogate model for bypassing Blade Element Momentum theory in wind turbine aerodynamic load estimation

Shubham Baisthakur, Breiffni Fitzgerald

Department of Civil, Structural and Environmental Engineering, School of Engineering, Trinity College Dublin, Ireland

ARTICLE INFO

Keywords:
- Physics-Informed Neural Network
- Surrogate modelling
- BEM aerodynamic model

ABSTRACT

This paper proposes the use of Artificial Neural Networks (ANNs), specifically Physics-Informed Neural Networks (PINNs), for dynamic surrogate modelling of wind turbines. PINNs offer the flexibility to model complex relationships while incorporating physics-based constraints, enabling accurate representation of wind turbine dynamics. In this paper, a PINN-based surrogate model is developed for the Blade Element Momentum (BEM) aerodynamic model used in state-of-the-art numerical wind turbine simulations. The PINN model replaces the time-consuming root-finding process in BEM with high-dimensional regression, significantly improving computational efficiency. The PINN model is trained using data generated from a numerical model of the IEA-15MW reference wind turbine, and its performance is compared against conventional data-driven Neural Network (DDNN) models. The proposed surrogate model provides more efficient and accurate evaluations of wind turbine responses compared with traditional surrogate modelling approaches. A significant computational advantage is obtained by using the developed surrogate models with a forty-fold speedup demonstrated compared to the BEM model. Replacing the BEM model with the PINN-based surrogate model for load computation in the numerical model used for dynamic analysis results in an overall reduction of 35% in computational time for a complete dynamic simulation. This is a substantial improvement in efficiency without sacrificing accuracy — the maximum Mean Absolute Error (MAE) values for the surrogate models are of the order of $10^{-2}$, which shows that the surrogate models can predict the angle of attack at any blade node with a discrepancy of less than 0.5°. The surrogate models significantly reduce computational time while maintaining high accuracy, making them a promising approach for simulating wind turbine dynamics, especially in fields such as reliability analysis or fatigue estimation where many simulations are necessary.

1. Introduction

As the adverse effects of fossil fuels on the Earth’s climate have become more evident, an increasing number of countries are transitioning towards sustainable sources of energy generation. This global push for sustainable development and energy security has led to significant investments in renewable energy, driven in part by the net-zero emission target set by the Paris Agreement, an accord signed by over 100 countries [1,2]. Among renewable energy sources, wind energy has gained particular attention and has seen substantial growth. To meet the rising demand, wind turbine installations must increase at an annual rate of 18% in this decade [3]. Offshore wind turbines, in particular, hold great potential as they benefit from higher wind velocities and a less turbulent environment, and they are expected to have a smaller impact on built urban environments [4].

Wind turbines have evolved significantly in both scale and technology over the past few decades [5]. These advances have significantly reduced the cost of wind energy generation, with the global-weighted average levelized cost of electricity (LCOE) from onshore and offshore wind declining by 13% and 9%, respectively [6]. This cost reduction can be mainly attributed to efficient controls, advanced technologies, and an improved understanding of wind turbine operations [7–14]. Despite these advancements, there is still an untapped potential to further reduce the cost of energy generation by optimizing wind turbine designs and ensuring consistent reliability during operation. However, the optimization and performance analysis require evaluating a large set of simulations for various combinations of operational parameters, which can quickly become computationally prohibitive. In this study, the authors aim to address this challenge by developing a surrogate model for predicting the aerodynamic loads on wind turbine blades. The use of this surrogate model will ensure a reduction in computational time, enabling optimization and reliability studies.

The current trend in wind energy is to design large standalone multi-megawatt wind turbines. These wind turbines are designed to...
have higher hub heights and larger rotor areas to maximize energy generation. The design of these large-scale wind turbines presents challenges. The increase in the size of the structural components (blades and towers) and the associated increase in flexibility combined with large dynamic loads from blade rotation, wind, and wave forces result in significant vibrations and structural fatigue [15]. Additionally, due to the increasing flexibility, the natural frequency of these multi-megawatt turbines is shifting towards the dominant frequency of sea waves, leading to an increased hydrodynamic loading [16]. This increases the risk of damage and reduces the overall lifespan of these turbines. With the upsaling of these wind turbines, blade tip deflection, web failure, and fatigue failure are becoming dominant modes of failure [17]. Fatigue loading, in particular, demands more attention, as a four-degree scaling relation between blade length and fatigue in root connections has been observed [17]. The risk of sustained vibrations, increased hydrodynamic loading, and higher levels of loads and deformations necessitate careful consideration of fatigue loading and structural integrity. The semi-probabilistic design guidelines based on the IEC 61400-1 standard cannot adequately address all uncertainties related to wind turbine operation, potentially leading to inconsistent reliability and infeasible design [18,19]. Especially the IEC-recommended guidelines for fatigue evaluation do not yield an entirely accurate estimation of the statistics. In addition, the design of serially produced wind turbines is typically based on a specific set of environmental parameters recommended by IEC for a particular site class, which may not be valid at all the installation locations. Evidently, this will lead to a modified load envelop and inconsistent reliability levels. To achieve consistent reliability levels, a large number of simulations are required to evaluate wind turbine responses for site-specific environmental parameters and compare them against the design values. However, this process is computationally intensive and requires access to high-performance computing facilities. Employing high-fidelity numerical models further compounds the computational challenges.

To overcome these issues, researchers have employed surrogate models, which offer a simplified representation of a complex system at a reduced computational cost. In the field of wind energy, Universal Kriging and Response Surface Method (RSM) based surrogate models are most commonly used. For example, Hu et al. [20] and Yang et al. [21] have used the Kriging surrogate model for reliability-based design optimization of wind turbine blade and tripod sub-structure for an offshore wind turbine, respectively, whereas Otkopokparoro and Sirimula [22] and Morató et al. [23] have presented the use of Kriging surrogate model for reliability analysis. Similar case studies of reliability analysis using the response surface method are widely presented in the literature [24–27]. Surrogate modelling has also been applied to site-specific load estimation of wind turbines, and a detailed comparison of commonly used surrogate modelling techniques in terms of accuracy and computational requirements was investigated by Dimitrov et al. [28].

Although RSM and Kriging based surrogate models are widely used, these are parametric regression techniques which assume a specific functional form. Also, these models assume a smooth response function and they may not work efficiently when the underlying variables have highly non-linear and complex behaviour. These conventional surrogate modelling techniques also cannot take into account the physical laws governing the wind turbine system. Whereas, the current developments in the field of Artificial Neural Networks (ANNs) provide flexibility to map the input–output relationship based on minimizing an explicitly defined function and satisfying the governing principles of the system [29,30]. Such constraints also minimize the admissible parameter space in the model discovery process and are highly efficient when only limited training data is available. To this end, ANNs are known to model highly non-linear relationships without assuming any underlying functional form and are capable of handling large data sets. Dimitrov et al. [28] also suggests that the ANNs or Machine Learning surrogate models could estimate the site-specific loads with higher accuracy, which is not considered in their study. Although conventional surrogate models are developed to reduce the computational costs, the development of these models still requires carrying out a large number of high-fidelity simulations for generating data sets for training these models, which leads to using more time and computational effort. Reducing the amount of data required for the training process can further improve the efficiency of these surrogate models.

In the context of state-of-the-art numerical wind turbine models like OpenFast and HAWC2, considerable computational effort is devoted to estimating aerodynamic loads for random inflow conditions using the Blade Element Momentum (BEM) theory. The BEM model’s computational expense arises from the time-consuming root-finding approach used to solve the governing equations. Previously, Fluck and Crawford [31] and Haghi and Crawford [32] have presented a stochastic solution for the unsteady aerodynamic loads using modified forms of polynomial chaos expansions. In these studies, the authors have bypassed the time-domain computation of wind turbine response by estimating the output distribution of the thrust and torque loads acting on the wind turbine rotor. These stochastic-based methods have been found to greatly reduce the computational time in predicting the output distribution and can be useful in optimization studies. However, both of these studies do not take into account the effect of controller algorithms which greatly impact the BEM output and wind turbine operation. Also, the analysis in these studies is performed assuming constant rotation speed, which limits their application. The model developed by Haghi and Crawford [32] aims to provide overall accurate statistics of the aerodynamic loading but point-to-point accuracy in the force estimation is questionable.

To address these challenges and improve computational efficiency, the primary objective of this research is to develop a Physics-Informed Neural Network (PINN) based surrogate model for the BEM aerodynamic model, which can introduce physics-based constraints in the training process. A weighing factor is introduced to optimally balance the contribution of the data-fitting objective and physics-informed criterion during the optimization process. The ANN based surrogate model aims to significantly enhance efficiency by replacing the root-finding process with high-dimensional regression. A numerical model of the IEA-15MW reference wind turbine is developed using Kane’s dynamics principle and benchmarked against OpenFast. Using the data generated from this model, a PINN model is trained to predict the angle of attack at each blade node as a function of inflow parameters. Unlike commonly used statistical mappings between input and output parameters, the PINN approach is developed to map inputs to outputs at each time step. This enables the surrogate model to be implemented in the existing wind turbine models for quicker response evaluations. The computational advantage gained by using the PINN model in place of the BEM aerodynamic model is investigated. A comparative analysis of the performance of PINN compared with the conventional data-driven NN model is presented to highlight its efficiency. By employing the PINN approach, this study aims to enable large-scale simulations for reliability analysis and site-specific investigations of wind turbines. The proposed surrogate model will facilitate a more efficient and accurate evaluation of wind turbine responses, ensuring consistent reliability levels and supporting wind turbine design optimization.

In this paper, a general overview of the BEM theory is presented in Section 2. The details of the numerical model of the IEA-15MW wind turbine model are presented in Section 3. This model is used to generate the training data. The methodology for developing the surrogate model is discussed in Section 4, and the numerical results and performance of the model is analysed in Section 5. This paper concludes in Section 6, summarizing the key findings of the study and the scope for future work in this area.
2. Blade Element Momentum (BEM) theory

Blade Element Momentum (BEM) theory is widely used for estimating aerodynamic loads on wind turbine blades and in many design and analysis applications. The theory assumes that the rotor operates in an ideal flow field, neglecting factors such as viscosity, compressibility, and three-dimensional effects. BEM theory combines the information from Blade element theory and Momentum theory to compute the angle between the resultant wind velocity and the chord line of the airfoil, commonly termed as the attack angle of an airfoil section, which is then used to compute the aerodynamic forces. Combining the information from momentum theory with blade element theory ensures that the forces and moments acting on each blade element can be determined while also considering the overall effect of the rotor or propeller on the flow. The momentum theory assumes a control volume section of the wind inflow and applies the principle of conservation of momentum to estimate the total thrust ($dT$) and torque ($dQ$) acting on an annular disk of radius $r$ and thickness $dr$ as shown in Fig. 1, and is given by Eqs. (1) and (2)

$$dT = \rho U^2 4\alpha (1 - a) \sigma r dr$$

$$dQ = \rho U d' (1 - a) \Omega r^2 \rho dr$$

where $\rho$ represents the mass density of inflow, $U$ is the inflow velocity far upstream of the actuator disk and $\Omega$ is the rotor speed. The terms $a$ and $d'$ represent the axial and tangential induction factors, which are the measure of induced linear and tangential wind velocity at the rotor plane due to the presence of the actuator disk. Momentum theory is useful for estimating the overall thrust produced but does not provide detailed information about the forces acting on individual blade elements. The blade element theory, on the other hand, divides the rotor or propeller into small blade elements and analyzes the forces and moments acting on each element. The blade element theory predicts the thrust ($dT$) and torque ($dQ$) as a function of the aerodynamic properties of the airfoil and the geometrical relationship between various forces acting on an airfoil section. The equation for thrust and torque at a distance $r$ from the blade root on an airfoil of chord length $c$ and radial length $dr$ using blade element momentum theory is given by Eqs. (3) and (4) where, $B$ is the number of blades in the rotor, $U_{rel}$ is the relative wind velocity which is the vector sum of wind velocity at the rotor and the wind velocity due to blade rotation, and $C_i$ and $C_d$ are the lift and drag coefficients of the airfoil section corresponding to the inflow angle $\phi$.

$$dT = \frac{1}{2} B \rho U^2 \rho r^2 \left( C_i \cos(\phi) + C_d \sin(\phi) \right) c dr$$

$$dQ = \frac{1}{2} B \rho U r^2 \left( C_i \sin(\phi) - C_d \cos(\phi) \right) c dr$$

The inflow angle is related to the angle of attack ($\alpha$) through Eq. (6), where $\theta_p$ is the sectional pitch angle, which is given by Eq. (5), where $\theta_p$ is the blade pitch angle which is a time-dependent parameter determined by the controller algorithm based on the region of operation of wind turbine and $\theta_T$ is the twist angle which is a geometrical property of the blade which varies along the blade length.

$$\theta_p = \theta_p + \theta_T$$

The geometrical relationship between inflow angle, wind velocity and induction factors is defined in Eq. (7), where $\lambda_r$ is the ratio of the tangential speed of blade tip to the inflow wind speed given by Eq. (8).

$$\phi = \theta_p + \lambda_r$$

$$\tan \psi = \frac{U(1 - a)}{\Omega r (1 + d')} = \frac{1 - a}{(1 + d') \lambda_r}$$

$$\lambda_r = \frac{\Omega r}{U}$$

The prediction of thrust and torque values obtained using the momentum and aerodynamic principles are equated to take into account both the global inflow characteristics and local aerodynamic properties. Combining this information gives the BEM theory’s prediction of the feasible lift coefficient ($C_l$) as a function of angle of attack ($\alpha$) for a constant tip speed ratio ($\lambda_r$) through Eq. (9). Also, from the principles of aerodynamics, the lift coefficient is an intrinsic attribute of the airfoil geometry and is expressed as a function of the angle of attack ($\alpha$) and the Reynolds number ($R_e$). Employing the aerodynamic characteristics of an airfoil ($C_l (R_e, \alpha)$) and BEM theory’s forecast of the lift coefficient as defined in Eq. (9), in conjunction with the geometric relationship outlined in Eq. (6), facilitates the iterative determination of the angle of attack ($\alpha$).

$$C_l = \frac{4 F}{\sigma r} \sin \phi \left( \frac{\cos \phi - \lambda_r \sin \phi}{\sin \phi + \lambda_r \cos \phi + \frac{C_d}{C_i} (\lambda_r \sin \phi - \cos \phi)} \right)$$
Fig. 2. Geometric representation of blade element theory.

The BEM equations, though conceptually simple, can be challenging to solve reliably and efficiently with high precision. Many solution approaches exist for numerically converging the angle of attack and axial and tangential induction factors but they generally suffer from a lack of robustness in some regions of the rotor blade design space or require significantly increased complexity to guarantee the convergence [33]. In this context, Ning et al. [34] has developed a robust method to solve the BEM equations with guaranteed convergence. The major contribution of this work was to parameterize the BEM equations through a single unknown variable, local inflow angle (\( \phi \)), by defining a residual function given by Eq. (10). Solving the residual equation by equating \( \zeta(\phi) = 0 \) gives the solution for the local inflow angle (\( \phi \)).

\[
\zeta(\phi) = \sin \phi - \frac{1}{1 - a} \frac{\cos \phi}{\lambda r} \cos \phi + a' (10)
\]

This approach allows to solve the BEM equations by using one-dimensional root-finding approaches, which are more efficient. The approach presented by Ning et al. [34] is implemented to compute the aerodynamic forces acting on the numerical model of the IEA-15MW wind turbine developed in this study. The solution for the residual equation is found using Brent’s approach [35]. A brief overview of the numerical model is presented in the next section.


The research undertaken in this study is performed using the specifications of the International Energy Agency (IEA) 15MW wind turbine, which is an IEC Class 1B direct-drive machine with a rotor diameter of 240 meters (m) and a hub height of 150 m. A representative model of the IEA 15MW wind turbine is shown in Fig. 3, reproduced from the technical report on the definition of the IEA 15-megawatt reference wind turbine [36]. This reference WT is chosen for this study as it represents the largest standalone WT model and is a leap ahead of current generation WTs [36]. The key parameters for this wind turbine are presented in Table 1. A multi-body dynamic model of this wind turbine is developed using Kane’s method [37]. Kane’s method reduces the labour needed to derive the equations of motion, and these equations are easier to model in a computer program than earlier classical approaches, such as the Euler–Lagrange method and D’Alembert’s principle. Kane’s approach is particularly useful in the analysis of multi-body dynamics, where different bodies are connected together to form a system. The equilibrium equations for a simple holonomic multi-body system using Kane’s approach is given by:

\[
F_r + F^* = 0
\]

where \( F_r \) stands for the generalized active forces and \( F^* \) represent the inertia force. These forces can be expressed in terms of the kinematic quantities as

\[
F_r = \sum_{i=1}^{n} E_{r X}^N \cdot F^X + E_{r \omega}^N \cdot M^N
\]

\[
F^* = -\sum_{i=1}^{n} E_{r \omega}^X (m^N \cdot E_{a X}^N) - E_{r \omega}^N \cdot E^{HN^N}
\]

where \( F^X \) is a force vector acting on the centre of mass of point \( X \) and \( M^N \) is the moment vector acting on the rigid body of \( N \). \( E_{r X}^N \) and \( E_{r \omega}^N \) are the partial linear and partial angular velocity of the point \( X \) and rigid body \( N \), respectively. \( E^{HN^N} \) is the time derivative of the angular momentum of the rigid body \( N \) about its centre of mass \( X \) in the inertial frame, given by the following equation

\[
E^{HN^N} = \dot{I}^N \cdot E_{a N}^N + E_{\omega N}^N \times I^N \cdot E_{\omega N}^N
\]

To describe the motion of the wind turbine, a total of 22 degrees of freedom (DOFs) are considered in this study. These DOFs represent:

1. Platform motion - 6DOF (3 translation + 3 rotation)
2. Tower deformation - 4 DOF (2 modes in fore-aft + 2 modes in side-to-side direction)
In this model, blades and tower are modelled as flexible members using the modal summation method. Coupled vibration modes for these elements are derived using BModes [38]. Various reference frames are defined to describe the motion of different components of the system and to define their orientation with respect to the other members. The final equation of the system is of the form

\[ M(q, \dot{q}) + f(q, \dot{q}, t) = 0 \]  \hspace{1cm} (15)

where \( M(q, t) \) is the inertia matrix and \( \dot{q} \) is the acceleration vector, whereas \( f(q, \dot{q}, t) \) is the force vector consisting of the external and restoring forces acting on the structure. This system of equations is solved using numerical methods. In this study we have used the fourth-order Runge–Kutta method. A complete derivation of these equations of motion is beyond the scope of this study and interested readers can refer to the work of Sarkar and Fitzgerald [39] for more insights.

The multi-body model developed in this study has been benchmarked against OpenFast and the results from the validation exercise are presented in Figs. 4 and 5. A good agreement between OpenFast and the numerical model derived in this study is observed. This model is further used to generate the data required for training the machine learning model. The main reason behind using the numerical model in this study instead of the OpenFast model is to compare the computational advantage of using the machine learning surrogate model, as this model can be easily integrated with the numerical model as compared to OpenFast.

### 4. Development of machine learning surrogate for BEM model

An ANN model takes an input feature vector \( x \), and defines a suitable hypothesis \( h_0(x) \) such that it can match the output \( y \). The NN hypothesis \( h_0(x) \) is designed using a series of interconnected NN layers, where a NN layer is characterized by the number of neurons and an activation function. A neuron, which is a building block of a NN architecture, is characterized by a specific weight (\( w \)) and a bias (\( b \)) value. A neuron receives the input \( x \) and produces the output as the weighted sum of input and bias value activated by an activation function. This mathematical operation for a single neuron is presented in Eq. (16), where \( f : R \rightarrow R \) is an activation function.

\[ h_0(x) = f(wx + b) \]  \hspace{1cm} (16)

The purpose of the activation function \( f(.) \) is to introduce non-linearity in the NN formulation, allowing the network to model more complex input–output relationships and shaping the decision boundaries. The selection of the activation function is an important step as it controls the network’s performance and learning capabilities; tanh, sigmoid, and ReLU are the most commonly used activation functions. In the context of a NN layer, in Eq. (16), the NN layer receives an input matrix \( x \) with the dimension of \( n_x \times n_y \) where \( n_x \) is the number of input features and \( n_y \) is the number of samples or observations and produces an output \( (h_0(x)) \) of dimensions \( n_y \times n_z \) where \( n_y \) denotes the number of nodes in the layer. The weight and bias matrix has dimensions of \( n_y \times n_x \) and \( n_y \times n_z \), respectively. Multiple NN layers can be arranged in a sequence such that the output of one layer serves as the input to another layer. For a general M-layer NN with a non-linear activation function \( f(.) \) and a parameter matrix \( \theta \) representing the learnable parameters, the NN hypothesis is represented in Eq. (17):

\[ h_M(x) = f_M(f_{M-1}(\theta_{M-1}, \ldots , f_2(\theta_2, f_1(\theta_1,x)))) \]  \hspace{1cm} (17)

Adding additional layers to the NN can help the model to capture more complex relationships between input and output, however, it is also more likely to overfit to the training data where the model becomes too specific to the training data and performs poorly on unseen data. To address this, a penalty or regularization term is often added to the objective function, discouraging overfitting and promoting generalizability. The training process of ANN revolves fundamentally around optimization. A suitable metric that aligns with the desired model behaviour and captures the errors between predicted and actual outputs must be defined, and then optimization algorithms iteratively adjust the weights and biases, optimizing the model’s performance based on the chosen loss function. Optimization algorithms such as stochastic gradient descent method (SGDM) [40] or adaptive moment estimation (Adam) [41] are commonly used to update the model’s parameters during the training process. Selecting a good set of input features and a suitable metric is not a trivial task.

#### 4.1. Training data generation

In this study, the objective is to develop a surrogate model of the BEM aerodynamic model, so naturally, the angle of attack is chosen for the training data generation. The angle of attack \( \theta \) is the angle between the wind vector and the horizontal plane on the blade.

### Table 1

Key parameters of the IEA-15MW wind turbine.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hub Height</td>
<td>150 m</td>
</tr>
<tr>
<td>Rotor Diameter</td>
<td>240 m</td>
</tr>
<tr>
<td>Cut-in wind speed</td>
<td>3.00 m/s</td>
</tr>
<tr>
<td>Rated wind speed</td>
<td>10.59 m/s</td>
</tr>
<tr>
<td>Cut-out wind speed</td>
<td>25.00 m/s</td>
</tr>
<tr>
<td>Minimum rotor speed</td>
<td>5.00 rpm</td>
</tr>
<tr>
<td>Maximum rotor speed</td>
<td>7.56 rpm</td>
</tr>
</tbody>
</table>

![Fig. 3. The IEA 15MW reference wind turbine.](image-url)
as the output parameter. From the description of the BEM model in Section 2, it is evident that the angle of attack is a function of inflow properties, control input, blade geometry, induction factors and blade vibrations. In this study, the surrogate model is trained to model the relationship between stochastic inflow parameters and control input with the angle of attack at a blade section for deterministic airfoil properties. The blade geometry is assumed to be constant. The controller inputs, specifically pitch angle and rotor speed, which are functions of the inflow parameters, are explicitly considered as the input features. These have been chosen as input features as this will make the developed
surrogate model applicable for different controller algorithms than the ones used while generating the training data. While selecting the input features for the surrogate model, only those parameters which can be measured during the wind turbine operation are selected. This will ensure that the developed model is useful in practical industry applications. The training data is generated using the numerical model presented in Section 3. In this study, the ROSCO controller is used while generating the training data [10]. In developing this surrogate model, the steady aerodynamic behaviour of the wind turbine is assumed and the effect of hub loss and tip loss is taken into account. The Pitt and Peters yaw correction model is used for yaw correction. Glauert’s correction with Buhl’s modification factor for high axial induction values is also considered [42]. The effect of varying Reynolds number on the aerodynamic behaviour of an airfoil is not considered during the development of the surrogate model. To this end, the wind velocity in $x$ and $y$ direction in the blade local coordinate system, rotor speed, blade pitch and generator azimuth angle are considered as the input features. The variation of the angle of attack (AOA) with respect to the input features for steady inflow conditions is presented in Fig. 6. For generating the training data, 50 mean wind speed data points with values ranging from cut-in to cut-out wind speed are generated using the Sobol sequence. Sobol sequence is used as it is known to have low discrepancy and better distribution properties. The scatter plot of generated wind velocities is presented in Fig. 7. Fig. 7 shows the efficiency of the Sobol sequence to effectively fill the parameter space. For these generated wind velocity values, random wind realizations are generated using the Turbsim module [43]. TurbSim is a stochastic, full-field, turbulent-wind simulator. It uses a statistical model to numerically simulate a time series of wind speed vectors in a two-dimensional vertical rectangular grid that is fixed in space. The wind speed time histories are generated for a turbulence intensity ranging from 5% to 10%, which is randomly selected for the generated wind speeds. This study assumes a normal turbulence model (NTM), defined in the IEC recommendations and the power law exponent of 0.12 is assumed to correspond to a relatively smooth and uniform surface [18]. Choosing a specific value of the power-law exponent does not limit the application of the developed surrogate model. For this set of parameters, wind speed realizations are simulated considering two random seed values for 100 s. The wind turbine response is evaluated for these conditions, and input and output features are recorded. For turbulent conditions, the response for the first 50 s is discarded to remove the transient effects. The response for the first random seed is selected for training, while the response corresponding to the second random seed is used for validation. The hold-out validation scheme is used during the model training.

4.2. Definition of objective function

In the training of an ANN surrogate model, the objective function plays a key role. An objective function is a metric chosen to quantify the agreement between the model output and the surrogate model prediction. The one-half squared error function is most commonly
used in regression-based surrogate models. Therefore, for the data-driven ANN surrogate, the one-half squared error function presented in Eq. (18), for a single observation, is used as the objective function.

\[ J_d(\theta) = \frac{1}{2} \| h_d(x) - y \|^2 \]  

(18)

To ensure generalizability and avoid overfitting, a regularization term is added to the objective function. The objective function for a set of input-output pairs is presented in Eq. (19), where \( g(\theta) \) represents the regularization term.

\[ J_d(\theta) = \frac{1}{m} \sum_{i=1}^{m} J(\theta; x^{(i)}, y^{(i)}) + \lambda g(\theta) \]  

(19)

In this study, the weight decay term is used as regularization as it discourages the neural network from relying heavily on individual weights and encourages it to use a more distributed representation of the input data. This leads to a more robust model and makes the model less sensitive to these small perturbations in input values. The final objective function used for training the data-driven ANN surrogate model is shown in Eq. (20), where \( \| \cdot \| \) represents the Frobenius norm of the weight matrix and \( n_L \) is the number of layers.

\[ J_d(\theta) = \frac{1}{m} \sum_{i=1}^{m} \| h_d(x_i) - y_i \|^2 + \lambda \sum_{j=1}^{n_L} \| W^{(j)} \|^2 \]  

(20)

Although the objective function presented in Eq. (20) can effectively capture the complex patterns in the data, it does not use any physical knowledge or previously known physical constraints of the problem at hand in the optimization process. This approach can struggle in regions not covered in the training data or where the available data is sparse and noisy. In these situations, physics-informed neural networks (PINNs) have an advantage over the purely data-driven approach as these models can leverage the prior knowledge of the system to guide the optimization process to learn the solutions which are physically more meaningful, reliable and indeed possible. Due to the additional constraints on the training data, PINNs can converge more quickly than their purely data-driven counterparts. In the context of the present problem, as presented in Section 2, axial and tangential induction factors provide meaningful information on the observed velocity at each blade section. These parameters are the latent variables of the system, as they cannot be explicitly measured. In this study, a PINN approach is used to define additional constraints using the latent system parameters such that the optimization process attains a solution where the residual term attains a minimum value such that the angle of attack, as well as induction factors, can have meaningful values. The objective function to enforce the physics-informed constraint is defined in Eq. (21):

\[ J_p(\theta) = \frac{1}{m} \sum_{i=1}^{m} R^2(h_p(x_i)) \]  

(21)

The additional constraint provided by Eq. (21) ensures that at the angle of attack predicted by the neural network hypothesis \( h_p(x) \), in addition to the minimization of mean square error between actual and predicted angle of attack, the residual term which determines the accuracy of the solution is also minimized. The term \( R^2(h_p(x)) \) ensures that the optimization process also minimizes the residual term defined Eq. (10). Although such physics-informed approaches appear to be straightforward and have yielded remarkable solutions for various applications, the effect of additional residual terms is poorly understood and, in several cases, can lead to unstable or erroneous predictions [44,45]. Therefore, to balance the interplay between data fit and physics-informed constraints, we introduce a weighing factor \( \beta \), which maintains a balance between the two loss terms in the final compositional form. The resulting optimization functional form for the PINN training is presented in Eq. (22).

\[ \arg \min_{\theta} \left( \beta J_d(\theta) + (1 - \beta) J_p(\theta) + \lambda \sum_{j=1}^{n_L} \| W^{(j)} \|^2 \right) \]  

(22)

For a data-driven neural network (DDNN) surrogate, the optimization functional form is presented in Eq. (23).

\[ \arg \min_{\theta} \left( J_d(\theta) + \lambda \sum_{j=1}^{n_L} \| W^{(j)} \|^2 \right) \]  

(23)

In this study, the above optimization problem is solved using the Adaptive Moment Estimation (Adam) optimizer. The optimum value of weighing parameter \( \beta \) is found in the hyperparameter tuning phase using a Bayesian optimization approach. The hyperparameter tuning process is presented in detail in the next section.

4.3. Hyperparameter tuning

In neural network (NN) models, hyperparameters represent key architectural settings that are determined prior to the model training.
process. Hyperparameter tuning is an important step in the development of an efficient NN, as these parameters directly impact the performance and generalization ability of the model. Effective hyperparameter tuning is pivotal for achieving an efficient and well-performing neural network. This process involves selecting the optimal combination of hyperparameter values, which can translate to heightened accuracy, quicker convergence, and improved generalization to novel, unseen data. In this study, we have used the Bayesian optimization technique in Matlab®. Bayesian optimization operates by constructing a probabilistic surrogate model of the objective function. This model facilitates informed decisions about generating new points within a predetermined range defined for the set of hyperparameters. This process involves an acquisition function that tactfully balances the exploration of uncharted regions and the exploitation of information at the current point.

Our study focused on a four-layer deep neural network with a $\tanh$ activation function. We performed the hyperparameter tuning across all fifty nodes representing the IEA-15MW wind turbine blade, where each node represents a unique airfoil section. Specifically, we optimized the number of neurons in each layer, the regularization factor, the learning rate, and the weighing factor. The ranges considered for these hyperparameters are detailed in Table 2.

Within this optimization framework, we adopted the “expected-improvement-plus” acquisition function. Notably, while the number of neurons in each layer, regularization factor, and learning rate was individually optimized for every blade section, our exploration led to particularly interesting insights regarding the weighing factor. The optimized value for the weighing factor corresponding to the fifty blade nodes is presented in Fig. 8, which shows that the weighing factor has attained a range of values.

Some of the blade nodes have the weighing factor value closer to 1, suggesting that the data-fitting constraint provides more valuable insights than the physics-informed constraint. This suggests that the PINN approach may not always lead to better performance in all circumstances. A careful consideration of the weighing factor is required for the specific problem under investigation. The final NN model for the optimized set of hyperparameter combinations is trained using the Adam optimizer. The results of training using the DDNN and PINN approach are presented in the next section.

5. Results and discussions

In this section, we present a comprehensive analysis of the machine learning surrogate models developed using the methodology presented in previous sections. This section presents a quantitative assessment of the performance of the machine learning surrogate in terms of computational advantage as well as the accuracy of the predictions to address the identified research questions. A total of nine nodes from the blade root to the tip section are selected and the surrogate models corresponding to these nodes are generated. These nodes are representative of the different operating conditions on a wind turbine blade and different shapes of airfoil sections. A physics-informed and data-driven NN surrogate for each of the selected blade nodes are developed and their performance is compared against each other. The NN-based surrogate is validated against the BEM aerodynamic model for both steady and turbulent conditions.

One of the most critical aspects when assessing the effectiveness of a surrogate model lies in quantifying the degree of error introduced when substituting the surrogate for the more complex, often more computationally expensive, model it represents. In this context, the selection of appropriate performance metrics becomes paramount to accurately gauge the surrogate’s fidelity. To this end, we have chosen two widely recognized metrics, namely the Mean Absolute Error (MAE) and Mean Square Error (MSE), to rigorously evaluate the surrogate model’s performance. MAE serves as a robust indicator of the average magnitude of the errors between the surrogate model’s predictions and the corresponding outputs from the BEM aerodynamic model. By focusing on the absolute values of these errors, MAE provides a clear and interpretable measure of the surrogate’s accuracy in approximating the BEM model’s behaviour. Additionally, we employ MSE, which quantifies the squared discrepancies between the surrogate’s predictions and the actual values of the angle of attack obtained from the BEM model. MSE, while sensitive to larger errors due to the squaring operation, offers valuable insights into the surrogate model’s precision and the extent to which it captures both minor and major variations in the data. The choice of these particular metrics is underpinned by their effectiveness in assessing the surrogate model’s ability to replicate the behaviour of the BEM aerodynamic model across a range of operating wind speeds. By employing MAE and MSE, we aim to provide a comprehensive understanding of the surrogate model's performance, ensuring that the evaluation is rigorous and informative. The magnitude of error at selected blade nodes is presented in Table 3 to highlight the accuracy of the model.
Table 3

<table>
<thead>
<tr>
<th>Blade node</th>
<th>Surrogate error PINN</th>
<th>Surrogate error DDNN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAE</td>
<td>RMSE</td>
</tr>
<tr>
<td>Node 5</td>
<td>0.0041</td>
<td>2.76e-05</td>
</tr>
<tr>
<td>Node 10</td>
<td>0.0017</td>
<td>4.623e-06</td>
</tr>
<tr>
<td>Node 15</td>
<td>0.0026</td>
<td>1.067e-05</td>
</tr>
<tr>
<td>Node 20</td>
<td>0.0015</td>
<td>3.331e-06</td>
</tr>
<tr>
<td>Node 25</td>
<td>0.0022</td>
<td>7.453e-06</td>
</tr>
<tr>
<td>Node 30</td>
<td>0.0028</td>
<td>1.302e-05</td>
</tr>
<tr>
<td>Node 35</td>
<td>0.0173</td>
<td>8.048e-04</td>
</tr>
<tr>
<td>Node 40</td>
<td>0.0038</td>
<td>3.757e-05</td>
</tr>
<tr>
<td>Node 45</td>
<td>0.0035</td>
<td>3.227e-05</td>
</tr>
</tbody>
</table>

The maximum MAE values for both the surrogates are of the order of $10^{-2}$, which shows that both the models can predict the angle of attack at any blade node with a discrepancy of less than 0.5°. This proves the accuracy of the surrogate model developed in this study. These error metrics are evaluated for both DDNN and PINN to compare their performance. From Table 3, it is observed that out of the nine nodes selected in this study, five nodes perform better under the PINN approach, while the remaining four models perform better under the PINN approach in terms of the data-fitting objective.

To enhance the assessment of model accuracy, Fig. 9 displays the values of the physics-based loss function. Remarkably, the figure reveals that five out of the nine selected nodes exhibit significantly lower loss values, while two additional nodes achieve comparable loss values. In aggregate, seven of the chosen nine nodes demonstrate improved accuracy concerning physics-based constraints under the Physics-Informed Neural Network (PINN) approach, with only two nodes surpassing the performance of the PINN method. Furthermore, the major benefit of the PINN surrogate is its ability to generalize and find the most optimal solution satisfying the underlying physics-informed objective function. This figure shows that PINN has a lesser error in these latent system parameters, which are equally important in the accurate prediction of the aerodynamic loads. In addition, the PINN model converges to the optimal solution in a lesser number of iterations, which leads to shorter model training time.

To further demonstrate the capability of the surrogate model, the time-history response of the PINN surrogate is compared against the BEM model for a range of wind speeds from cut-in to rated wind speeds for both steady wind and turbulent wind with 10% turbulence intensity. Fig. 10 demonstrates the capability of the developed surrogate to accurately predict the angle of attack at each time step over the range of wind speeds in both steady and turbulent conditions. This capability of the surrogate model allows it to be integrated with the state-of-the-art wind turbine models used for the dynamic analysis of the wind turbine structure. In addition, the coefficient of determination ($R^2$) for the selected blade node is also presented in Fig. 10. The $R^2$ value is a statistical measure of the goodness of fit of the regression model. Here, the $R^2$ value is more than 0.99, which further demonstrates the accuracy of the surrogate model. Fig. 10 also presents the distribution of error in the predicted and target values against the target values. A uniform distribution of error is observed over the range of angle of attacks, which shows that the developed surrogate model does not suffer from heteroscedasticity commonly discussed in the literature. The estimated values of aerodynamic forces using BEM theory, PINN surrogate model and DDNN surrogate are compared in Fig. 11. For both the surrogate models, the percentage error in predicting aerodynamic forces is less than 5% in all cases.

The accuracy of the surrogate models and their capability in predicting the angle-of-attack and aerodynamic forces across different wind speeds and operating regions is clearly demonstrated in Figs. 10 and 11. The other major aim of the surrogate models is to reduce the computational time involved in evaluating the BEM equations. The NN surrogates developed in this study are capable of predicting more than 80000 observations/sec, which is more than 40 times quicker as compared to the conventional way of predicting the angle-of-attack through iterative root finding methods. Especially in this study, the comparison is made against Ning’s method [33] of solving BEM equations, which is the most robust method and is also implemented in OpenFast. Since the surrogate model is capable of estimating the angle-of-attack at every time step, the model is integrated with the numerical model of the wind turbine presented in Section 3. The use of our surrogate model in place of the traditional aerodynamic BEM model has yielded a substantial reduction in the overall simulation time of the numerical model, equating to an approximate 35% reduction in computational time. The reduction in computational time is calculated by comparing the execution time of the numerical model required to predict the 600-second response to a turbulent wind inflow using the conventional root-finding method and surrogate model to solve the BEM equations. The reduced computational time allows for running large sets of simulations for e.g. reliability estimation, fatigue estimation or optimization problems. The proposed surrogate model will especially benefit the site-suitability studies where the numerical model has to be evaluated for the load envelope at multiple sites to select the most optimum installation location. All the simulations in this study are performed on a system with an 8-core Intel Xeon CPU with a clock speed of 3.8 GHz using 32 GB RAM and running on Microsoft Windows 10 Pro.
6. Conclusion

In this study, we present a machine learning-based surrogate approach aimed at estimating aerodynamic forces on wind turbine blades under stochastic inflow conditions. This methodology showcases the potential of combining data-driven and physics-informed techniques to enhance the robustness and efficiency of neural network models. To train the surrogate models, we employ a multi-body dynamic numerical model of the IEA-15MW wind turbine, which is benchmarked against OpenFast. This benchmarking process ensures the accuracy and reliability of the training data. For sampling mean wind speeds, the Sobol sequence is utilized, while TurbSim is used to generate turbulent wind speed realizations. Feed-forward neural network algorithm is used for generating surrogate models, capable of delivering precise estimations.

This study extends the scope of surrogate modelling by considering physics-informed constraints along with data-fit objective functions. This approach leads to accelerated convergence and the development of a robust neural network model. A weighing factor which balances the relative importance of these two objectives is introduced and is optimized using the Bayesian optimization approach. A quantitative assessment of the surrogate models’ accuracy showcases their ability to closely approximate the BEM aerodynamic model. Notably, the physics-informed surrogate demonstrates even closer alignment with the original BEM model, offering an advantage over the data-driven methodologies.

One of the defining features of this proposed surrogate model is its ability to predict the angle of attack and, in turn, the aerodynamic forces at every time step, as opposed to commonly used statistical mapping in the conventional surrogate modelling techniques. This capability of the surrogate model allows for the integration of these models with numerical models used for the dynamic analysis of wind turbines. The estimation of load time–history response allows these surrogates to be used for fatigue analysis, which is not possible with the surrogate models developed for statistical mapping of input and output.
domains as highlighted by Haghi and Crawford [32]. In addition, a significant computational advantage is obtained by using these surrogate models. In standalone aerodynamic load computations, a remarkable forty-fold speedup as compared to the conventional root-finding approach to solve BEM equations is observed. Replacing the BEM model with the neural network-based surrogate for load computation in the numerical model used for the dynamic analysis in this study leads to an overall reduction of 35% in the computational time required for a complete dynamic simulation. This computational speedup empowers large-scale simulations necessary for uncertainty quantification, reliability analysis, fatigue calculations, and optimization tasks at a significantly reduced computational cost. In terms of computational efficiency, both data-driven and physics-informed surrogates have a similar performance.

While these surrogates are designed for airfoil shapes specific to the IEA-15MW wind turbine, the methodology itself is adaptable to any wind turbine design. One of the limitations of this work is that these surrogates do not explicitly consider the effect of Reynold's number or wind shear on the angle of attack or aerodynamic forces, as the aim of this work was to consider only physically measurable parameters as the input features. This choice of input features opens the possibility of enriching the surrogate models with the experimental data to consider the realistic behaviour observed in operating wind turbines.

In our investigation involving nine different blade nodes and airfoil shapes, we observed a mixed performance between data-driven and physics-informed neural networks. Data-driven neural networks exhibited superior performance in only two of these cases, prompting us to scrutinize the underlying factors contributing to this discrepancy. Upon closer examination, it becomes evident that specific conditions, airfoil shapes, and flow characteristics play a decisive role in determining the network’s effectiveness. This outcome underscores the importance of proper selection between data-driven and physics-informed neural networks for surrogate modelling in aerodynamics. Moreover, it reaffirms that the choice between these approaches should not be made with a one-size-fits-all approach but rather by assessing the problem’s unique demands. While our findings highlight the versatility of data-driven networks in certain scenarios, it is imperative to recognize the enduring value of physics-informed networks within the wider context of aerodynamics and surrogate modelling. Moving forward, further research can delve into identifying the conditions conducive to data-driven network excellence and refining physics-informed networks to address the limitations elucidated in this study. However, irrespective of the distinction between data-driven and physics-informed, the neural network-based surrogate approach presented in this study offers an advantage over the conventional surrogate modelling techniques both in terms of computational time and accuracy. Another major advantage is these model’s ability to predict the response at every time-step during the dynamic analysis enabling their integration into the state-of-the-art numerical models of wind turbines.

The methodology proposed in this study lays the foundation for expanding our research trajectory into the domain of wind turbine meta-modelling. In particular, the development of a surrogate model to predict the response of a wind turbine blade subjected to the aerodynamic loads can completely bypass the computationally inefficient numerical models with a more efficient data-driven metamodel. Furthermore, this approach holds promise for applications in health monitoring and the creation of digital twins for operational wind turbines, utilizing sensor and SCADA data. In addition, the PINN surrogate model-based approach for bypassing BEM load estimation, such as is proposed in this work, must be considered in practical terms with concepts from areas such as high-performance materials and advanced manufacturing technologies [46–48]. This is beyond the scope of the current work presented in this paper but may form the basis of future studies and may be necessary for practical implementation in the field.

CRediT authorship contribution statement

Shubham Baisthakur: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Breiffni Fitzgerald: Writing – review & editing, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.
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

The data presented in this study are available on request from the corresponding author. The data are not publicly available because it also forms part of an ongoing study.

Acknowledgements

This work is supported by Science Foundation Ireland (SFI) grant no. 20/FFP-P/8702 and Sustainable Energy Authority of Ireland (SEAI) grant no. 19/RDD/511.

References