RESEARCH ARTICLE

WILEY

Surrogate models for parameterized representation of wake-induced loads in wind farms

Nikolay Dimitrov [iD]

Department of Wind Energy, Technical University of Denmark, Roskilde, Denmark

Correspondence
Nikolay Dimitrov, Department of Wind Energy, Technical University of Denmark, Frederiksborgvej 399, Roskilde 4000, Denmark.
Email: nkdi@dtu.dk

Abstract

A procedure for mapping wake-induced load predictions computed with the dynamic wake meandering model to a computationally efficient surrogate model approximation is defined and demonstrated. Using the mapping function, the load variation can be efficiently estimated for a wind farm with arbitrary layout. The resulting load assessment procedure provides continuous, differentiable output with known analytical derivatives and can be used for applications such as wind turbine layout optimization, estimation of turbine lifetime, and uncertainty analysis.

KEYWORDS

layout optimization, loads, neural networks, polynomial chaos, probabilistic, surrogate, wake, wind farm

1 | INTRODUCTION

The wake-induced load effects experienced by wind turbines within a wind farm are an important design factor, which can impose restrictions on the wind farm layout or the turbine design choices. Therefore, wake-induced load analysis is a central part of wind farm planning and design. One of the pioneering approaches was the so-called Sten Frandsen model,[1] which takes wake effects into account by assigning an equivalent ambient turbulence level, which should cause the same fatigue load accumulation as the actual wake-induced turbulence and deficit. The Sten Frandsen approach is widely popular due to its simplicity, but it typically results in conservative load estimates.[2] A more advanced method is the dynamic wake meandering (DWM) model,[3] which provides an engineering approach to including a wake deficit in the turbulence box used for aeroelastic load simulations. The capability of the DWM model to predict wake-induced load effects has been validated for specific turbines in a wind farm,[2,4-6] as well as on turbine prototype test sites.[7] In the latter, the DWM model was shown to have superior performance than other engineering models. A table-lookup-based approach for mapping wind farm power output based on the DWM model has been demonstrated by Keck and Undheim,[8] while a procedure for mapping the load variation within a wind farm is shown in Galinos et al.[9]

The aim of the present study is to define and demonstrate a simple procedure that maps the outputs of the DWM model in a computationally efficient surrogate model approximation so that the load variation can be efficiently estimated for a wind farm with arbitrary layout. The resulting quick load assessment provides continuous, differentiable output and can be used for applications such as wind turbine layout optimization, estimation of turbine lifetime, and structural reliability analysis. The main hypothesis is that the wake-induced effects on loads can be described by means of variables related to the wind farm geometry and the ambient conditions and that a surrogate model calibrated by machine learning from a large load simulation database can be used to efficiently quantify such a relation. The parameterization of wake-induced load effects is obtained under the assumptions underlying the DWM model, ie, that the wake deficit behaves as a passive tracer and follows the transverse wind fluctuations and that the ambient turbulence causing the meandering can be described by a Gaussian random turbulence model such as the Mann model.[10]

2 | FORMULATION

2.1 | Descriptive variables for load variation within a wind farm

The load variation within an offshore wind farm is primarily a product of wake-induced flow disturbances. The form and severity of the wake-induced turbulence and deficit depend on multiple factors, which can be grouped into three classes:

(a) ambient conditions (wind speed, ambient turbulence, atmospheric stability, wave height, and others);

(b) wind turbine operational regime (rotor thrust, rotational speed, and power set point); and

(c) relative position of the wake source(s) with respect to the disturbed turbine.

Most of the aspects of the wind turbine operational regime are a function of the ambient conditions, and the relationship is an integral part of the high-fidelity load assessment tool, which normally utilizes information about the turbine control strategy. Thus, any variables describing the operating regime are in general dependent variables, and their influence can, through the load model, be expressed in terms of the environmental variables. An exception is an eventual power curtailment (eg, to limit or optimize the power output of large wind farms by derating individual turbines), which is controlled externally and therefore should be considered as an independent variable. Another potential source of load variation over an offshore wind farm that should be considered is the difference in the support structures (eg, due to depth and seabed conditions) and other possible variations of the structural properties of the turbines. Let us denote a load quantity of interest (eg, damage-equivalent loads) by $S(\mathbf{X})$, where $\mathbf{X}$ is a set of input variables. On the basis of the grouping defined above, the elements of $\mathbf{X}$ will be

$$\mathbf{X} = [\mathbf{X}_{amb}, \mathbf{X}_{pos}, \mathbf{X}_{farm}], \tag{1}$$

where the *amb*, *pos*, and *farm* subscripts represent variables related to the ambient conditions, turbine position, and wake-induced effects, respectively. The precise choice of individual variables may depend on the farm configuration and the purpose of the study—eg, considering wave conditions is an imperative when assessing loads on offshore foundations, however, it may be irrelevant if the component of interest is in the rotor-nacelle assembly, which is typically not affected by wave-induced loads. If the intention is optimizing wind farm control, an additional group of variables, $\mathbf{X}_{op}$, would be needed to account for the control actions such as derating or wake steering. In order to allow generalizing the model results for arbitrary wind farm layouts and turbine positions, the variables in $\mathbf{X}_{farm}$ should not reflect a specific turbine location or a specific wind farm but should instead define parameterized conditions in terms of relative distances to neighboring turbines and orientations with respect to the wind direction. The farm layout and the presence of regular patterns can nevertheless affect the choice of variables describing wake-induced effects. The variables defined below are suitable for farms with predominantly regular layout, and additional variables may be necessary for other cases. For the present study, it is chosen to work with the following variable set:

$$
\begin{aligned}
\mathbf{X}_{amb} &= [\bar{u}, \sigma_u, \alpha, H_s, T_p, \Delta] \\
&\quad (\text{mean wind speed, turbulence, wind shear,} \\
&\quad \text{significant wave height, wave peak period, and wind -wave misalignment}) \\
\mathbf{X}_{pos} &= Z_w (\text{water depth}), \text{ and} \\
\mathbf{X}_{farm} &= [R_D, \theta, N_{rows}] (\text{row spacing, wake incidence angle, and number of disturbing turbines, see Figure 6}).
\end{aligned}
$$

The choice of the environmental variables given above reflects the most relevant factors, which could influence loads on both the rotor-nacelle assembly and the foundations—and since the foundations are taken into consideration, the water depth is also considered as part of the position-specific variables. For maintaining simplicity, it is chosen to omit the independent variables related to farm operation and control ($\mathbf{X}_{op}$), but they can in principle be included without changing the modelling arrangements presented here. The wake-related variable choice needs slightly more detailed explanation, which follows.

2.2 | Wake meandering model parameters

From the location of the rotor plane of a wind turbine aligned with the ambient wind direction, the wake from a nearby turbine can be seen as a roughly circular area with reduced wind speed (wake deficit) and modified turbulence (Figure 1). The wake position and size are defined by the location of the wake centre (y_c, z_c) and by the wake diameter D_w, which is the average diameter of the area with wind deficit A_w, where $D_w = 2\sqrt{A_w}$. The wake diameter just behind the turbine is denoted as D_0 and continuously increases downwind due to the expansion of the wake.
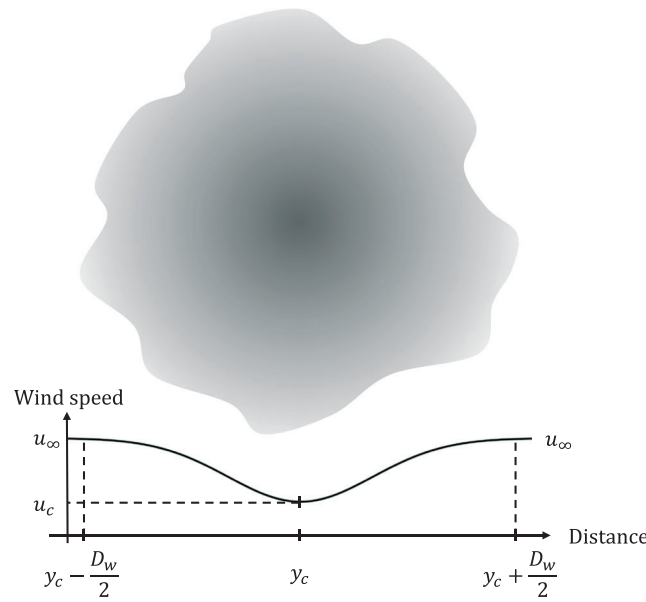
FIGURE 1 Example wake deficit with width D_w, centre position y_c, centreline velocity u_c, and free wind velocity u_∞

The wake-induced changes in loads will depend on the relative location and size of the wake deficit (what part of the rotor is affected by it) and by the turbulence characteristics within the wake-affected volume. The assumption used in the DWM model definition is that the wake is advected downwind and its centre position follows the fluctuations of the transverse wind.[3] In the same study, the change in the wake centre position y_c (Figure 1) between time step t and $t+\Delta t$ is defined as

$$y_c^{(t+\Delta t)} = y_c^{(t)} + v_{char}(u(T-t), y_c^{(t)}, z_c^{(t)}), \Delta t \tag{2}$$

where v_{char} is a characteristic transverse wind velocity of the wake deficit at time t, T is the total length of the wind field simulation in seconds, $u(t)$ is the longitudinal, free wind velocity determining the speed of downstream advection, and y_c and z_c are, respectively, the wake centre horizontal and vertical positions. The characteristic transverse velocity v_{char} is defined as the two-dimensional integral of the transverse wind field velocity v_c over the area of the wake deficit. On the basis of Equation (2), y_c can be considered as a random variable, which is proportional to the sum of the jointly Gaussian distributed, zero-mean random variables $v_{char}(t)$. It follows that y_c is also Gaussian, with zero mean and with variance that can be computed by making use of the simple identity $\sigma^2(ax_i+bx_j)=a^2\sigma^2(x_i)+b^2\sigma^2(x_j)+2ab\text{Cov}(x_i,x_j)$ where $\sigma^2(\cdot)$ and $\text{Cov}(\cdot)$ denote variance and covariance, respectively. For any distance downwind, the variance of y_c is computed as follows:

$$\sigma_{y_c}^2(\sigma_{v_{char}}^2, t) = \sum_{i=1}^{N=t/\Delta t} \left\{ \Delta t^2 \left\{ \sigma_{v_{char}}^2 + 2 \sum_{j=1}^{i-1} \text{Cov}(v_{char,j}, v_{char,i}) \right\} \right\}, \tag{3}$$

where Δt is the simulation time step and $\text{Cov}(v_{char,j}, v_{char,i})$ denotes the covariance between the characteristic transverse wind speed components at time steps i and j.

From the definition of v_{char} as the area average of point velocities, its covariance properties will be a low-pass filter function of the point variance of the transverse wind speed, σ_v^2. Knowing the low-pass filter properties and the point turbulence spectrum (eg, Mann turbulence model[10] with parameters L, Γ, and $\alpha\varepsilon^{-2/3}$), the covariance of v_{char} as a function of spatial separation can be determined from the autocorrelation of the v component of wind speed by inverse Fourier transforming the v component of the Mann spectrum:

$$R_v(\Delta v) = \frac{1}{\bar{S}_{vv}} \int_{-\infty}^{\infty} S_{vv}(k_1)e^{-i\Delta v k_1} dk_1, \tag{4}$$

where $S_{vv}(k_1)$ is the v-component spectrum, k_1 is the wave number related to the frequency f as $k_1=2\pi f/u$, and $\bar{S}_{vv}$ denotes the mean of the spectrum, serving for normalization. In Equation (4), $S_{vv}(k_1)$ represents the point spectrum, which is computed by numerical integration of the Mann spectral tensor[10] over wave numbers (k_2,k_3) representing the coordinate axes perpendicular to the wind direction. Low-pass filtering is applied by restricting the integration limits for k_2 and k_3 to wave numbers below the cutoff limit $k_{max}=\pi/D_0$. The choice of cutoff limit is based on Larsen

et al.[3] Computing $S_{vv}(k_1)$ with this approach and using it in Equation (4) result in the autocorrelation function for v_{char}. For any given time t and v-component variance σ_v^2, the probability distribution of the wake centre position is thus given by

$$p_{y_c}(y_c, \sigma_v^2, t) = \phi\left(\frac{y_c - y_0}{\sigma_{y_c}(\sigma_{v_{char}}^2, t)}\right), \tag{5}$$

where $y_0 = y_c(t=0)$ denotes the initial wake position at $t=0$s and $\phi(\cdot)$ denotes the standard normal distribution. The probability distribution of the wake centre position $p_{y_c}(y_c, \sigma_v^2, t)$ for initial wake position $y_0=0$ is illustrated in Figure 2: the uncertainty increases significantly with the distance from the source turbine.

For the purpose of computing wake incidence probabilities, a wake condition is considered as a situation where any part of the wake deficit of an upwind turbine coincides with any part of the rotor of a downwind turbine—meaning that the range of downwind locations that are potentially affected by wakes is approximately one wake diameter larger than the range of wake centre positions. For any given point with coordinates (t,y) downwind of a wake source, the probability $p_w(t,y)$ of any part of a wake disturbance being incident on the point will depend on the initial wake position y_0, which can be defined as a random variable with probability distribution $p_{y_0}(y, \sigma_{y_0}^2)$, the wake position variance $\sigma_{y_c}^2(\sigma_v^2, t)$, and on the wake diameter $D_w(t)$ where the latter changes with time due to wake expansion. This probability can be evaluated using a double convolution:

$$p_w(t, y) = \int_{y_2=-\infty}^{y_2=\infty} p_{y_0}(y_2, \sigma_{y_0}^2) \int_{y_1=y-D_w(t)/2}^{y_1=y+D_w(t)/2} p_{y_c}(y_1, \sigma_v^2, t)dy_2dy_1. \tag{6}$$

In the above, y_0 is deterministic for a wake source with fixed position such as a single upwind turbine. As a result, only the inner integral with respect to y_1 needs to be evaluated. Figure 3 shows plots of the computed wake incidence probabilities for $y_0=0$, downwind distance seven rotor
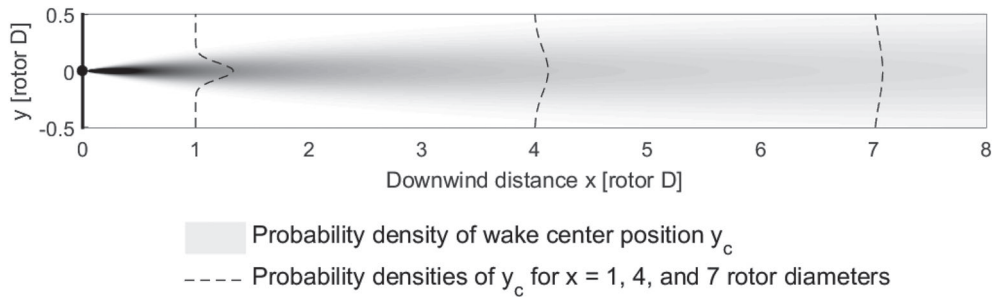
FIGURE 2   Probability distribution of wake centreline, assuming initial wake position $y(t=0)=0$
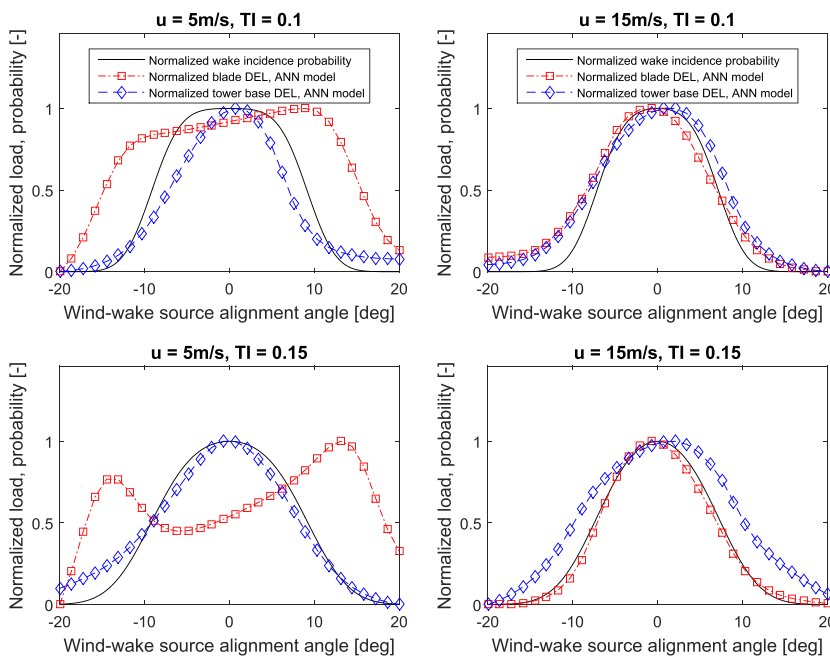
FIGURE 3   Comparison of wake incidence probability and normalized damage-equivalent loads as functions of the wind-wake source misalignment angle (the difference between the mean wind direction and the upwind turbine azimuth relative to the downwind turbine). Data shown correspond to downwind distance of seven rotor diameters [Colour figure can be viewed at wileyonlinelibrary.com]

diameters, and several combinations of mean wind speed $\bar{u}$ and turbulence intensity $TI = \sigma_u/\bar{u}$. The probability densities are normalized so that the peak density equals to 1. The normalized damage-equivalent fatigue load (DEL) predictions from an artificial neural network (ANN) surrogate model (see Section 2.3.2) for two load channels are overlaid on each probability plot. The curves show the load excess, ie, the difference between the load experienced under wake effects and under free wind conditions, normalized so that the maximum excess equals 1. The load behaviour curve obtained in this way has a strong resemblance with the probability curve, showing that indeed, the wake effects can drive load increase. The best agreement is obtained for the tower base loads, while the blade root flapwise loads show a double-peak behaviour at low wind speed, which is due to partial wake situations. In a partial wake situation, the wake deficit affects only a part of the rotor area, meaning that the blades experience cyclic fluctuations in the incoming wind speed. This in terms affects the accumulated fatigue loads.

The expression for the probability distribution of the wake centre position can be used to estimate what is the range of alignment angles between the free wind direction and a wake source turbine, which includes a possibility for experiencing wake conditions. This is illustrated in Figure 4 by showing the 95% confidence intervals for the wake centres as well as for the wake edges (the boundaries of the areas affected by deficit) for several turbines in a cluster. The wake widths shown in Figure 4 take into account wake expansion (initial expansion and downwind expansion) using the dependencies defined in Larsen et al[3] and a C_t curve for a pitch-controlled wind turbine with 80-m rotor diameter. The initial wake position is considered as deterministic with $y_0=0$. The angular ranges that represent the 95% confidence interval width for encountering wake situations are shown in Table 1 as function of wind speed, turbulence intensity, and downwind distance in rotor diameters. As expected, the wake-affected range is larger for higher turbulence and lower wind speeds as both factors increase the amount of meandering. It seems that for most offshore wind farms where the closest spacing is at, eg, five to seven rotor diameters, the wake-affected directions are confined to ±20 ° relative angles and decreasing with lower turbulence and larger spacing.

With the above, it is clear that the wake-induced effects are a function of the turbine spacing and the angle between wind turbine rows and the wind direction and that these variables are also dependent on wind speed and turbulence. With this variable space, it is possible to characterize the wake-induced effects caused by single turbines. However, the DWM model allows the superposition of wake deficits from multiple source
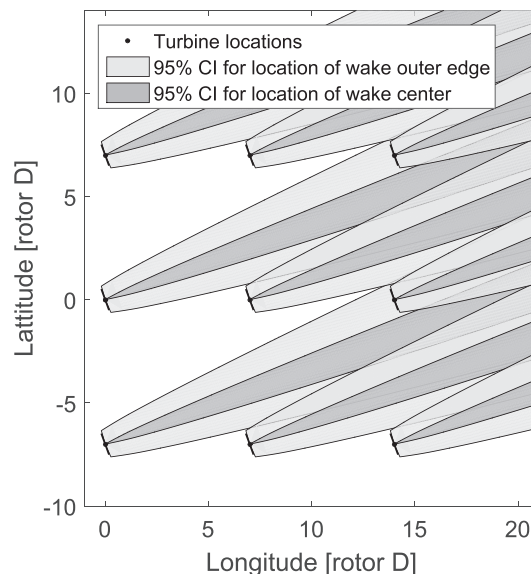
FIGURE 4 Interaction of wakes including wake expansion and misalignment between wind directions and row alignment

TABLE 1 Range of wind-source turbine alignment angles that represent the 95% confidence interval for experiencing wake situation in a load simulation using the dynamic wake meandering model

|	$\bar{u}$ = 5 m/s Distance in Rotor D				$\bar{u}$ = 10 m/s Distance in Rotor D				$\bar{u}$ = 15 m/s Distance in Rotor D			
TI	5D	10D	15D	20D	5D	10D	15D	20D	5D	10D	15D	20D
0.05	27.4	19.0	16.0	14.4	25.8	17.9	15.1	13.5	22.4	15.6	13.1	11.8
0.1	32.0	22.9	19.3	17.3	30.4	21.7	18.3	16.5	27.1	19.4	16.4	14.7
0.15	36.7	26.6	22.5	20.2	34.9	25.4	21.5	19.3	31.6	23.1	19.6	17.6
0.2	41.2	30.3	25.7	23.1	39.3	29.0	24.7	22.2	36.1	26.7	22.8	20.5

Note. The range widths are given in degrees.

turbines in a row—a so-called deficit summation. This is relevant especially at higher wind speeds when the thrust coefficient is relatively low and thus, the deficits from individual turbines are not very pronounced.[2] The impact on wind turbine loads from deficit summation from an increasing number of wind turbines in a row is shown in Figure 5, for mean wind speed 15 m/s and spacing of seven rotor diameters. The change in loads is most pronounced when one or two consequent wake sources are added; the effect seems to saturate for four or more turbines in a row. To summarize, the wake-induced effects on loads simulated using the DWM model can be approximately characterized using the dependence of wake properties on three variables: row spacing, wake incidence angle, and number of disturbing turbines, as shown in Figure 6.

2.3 | Surrogate models

The purpose of using a surrogate model in the present study is to obtain an efficient function, which provides continuous mapping between input variables and outputs such as loads and power. Calibrating this function against higher fidelity simulations allows replacing the computationally expensive models with a quick and useful approximation. In Dimitrov et al,[11] several wind turbine load surrogate model choices are compared, and the polynomial chaos expansion (PCE) is recommended as the potentially best choice out of the models considered. Further, in Schröder et al,[12] ANNs are shown to match or exceed the performance of the PCE on a similar problem. The present work employs surrogate models based on PCE and ANN and compares the surrogate model performance achieved with the two approaches.

2.3.1 | Polynomial chaos expansion

Polynomial chaos expansion is a popular method for approximating a stochastic function of multiple random variables using an orthogonal polynomial basis. For a stochastic function $S(\mathbf{X})$ of a finite number of random variables $\mathbf{X}=[X_1,X_2,...,X_M]$, a PCE will theoretically represent an infinite
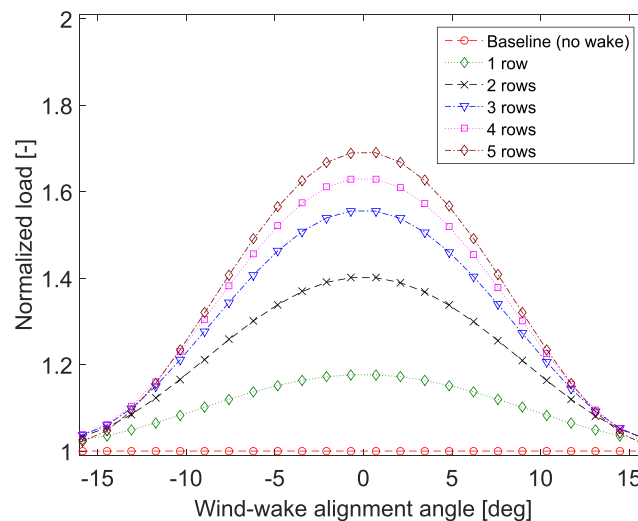
FIGURE 5   Illustration of the effect of wake-induced turbulence on blade root flapwise loads [Colour figure can be viewed at wileyonlinelibrary. com]
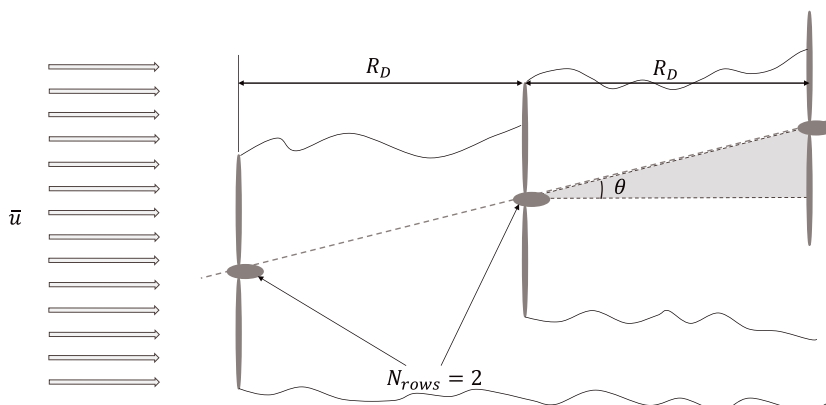
FIGURE 6   Illustration of a row of a turbine in a wind farm with the variables characterizing the wake interaction: R_D is the row spacing in rotor diameters, θ is the angle between the wind direction and the row alignment direction, and N_{rows} is the number of turbines upwind from the turbine under consideration [Colour figure can be viewed at wileyonlinelibrary.com]

sum of orthogonal polynomials. In a practical application, the sum is truncated to a finite number of terms depending on the number of variables M and with a maximum degree p:

$$\tilde{S}(\boldsymbol{\xi}) = \sum_{j=1}^{N_p} S_j \Psi_{\alpha,j}(\boldsymbol{\xi}), \tag{7}$$

where N_p is the total number of terms, $S_j \in \mathbf{S} = [S_1,...,S_{Np}]$ are unknown coefficients that need to be determined, $\boldsymbol{\xi} = [\xi_1,...\xi_M]$ are functions of $\mathbf{X}$, and $\Psi_{\alpha,j}(\boldsymbol{\xi})$ are multivariate orthogonal polynomials of dimension M.

In the classical PCE definition (see, eg, Ghanem and Spanos[13]), the input random variables are normally distributed or transformed to normal variables, which means that the Hermite polynomials are a suitable basis as they are orthogonal with respect to the normal probability density. Other choices of polynomial basis are also possible, which turns the problem into the so-called Wiener-Askey or generalized chaos.[14] For the present application, the Legendre polynomials are considered as the most convenient as they are defined on an interval $[-1,1]$, which can easily be mapped on the cumulative distribution functions of the random variables $\mathbf{X}$ using a simple transformation:

$$\xi_i = 2F(X_i) - 1, \tag{8}$$

where $F(\cdot)$ denotes the cumulative distribution function.

The univariate Legendre polynomial family for dimension i and with maximum order p is here denoted by

$$P_{\alpha_i}(\xi_i), \qquad \text{where} \quad i = 1, \ ..., M, \quad (\alpha_i \geq 0) \in \mathbb{N}, \quad \sum_{i=1}^{M} \alpha_i \leq p. \tag{9}$$

In the above, α_i denotes the order of the individual univariate polynomials, p the order of the multivariate polynomial, and $\mathbb{N}$ is the integer set. On the basis of this definition, the multivariate polynomials of dimension M are then defined as

$$\Psi_\alpha = \prod_{i=1}^{M} P_{\alpha_i}(\xi_i). \tag{10}$$

The total number of polynomials of this type is[15]

$$N_p = \frac{(M+p)!}{M!p!}. \tag{11}$$

The most straightforward way of determining $\mathbf{S}$ from a set of observations $g(\boldsymbol{\xi})$ is minimizing the variance of the residual $\varepsilon = g(\boldsymbol{\xi}) - \tilde{S}(\boldsymbol{\xi})$ using a least-squares regression approach:

$$\mathbf{S} = \text{Argmin} \frac{1}{N_e} \sum_{i=1}^{N_e} \left\{ g(\boldsymbol{\xi}^{(i)}) - \sum_{j=0}^{N_p-1} S_j \Psi_{\alpha,j}(\boldsymbol{\xi}^{(i)}) \right\}^2, \tag{12}$$

where N_e is the number of sampling points in the experimental design. For this purpose, a design experiment has to be set up, and the so-called design matrix $\boldsymbol{\Psi}$ needs to be constructed:

$$\Psi_{ij} = \Psi_{\alpha,j}(\boldsymbol{\xi}^{(i)}), \qquad i = 1, \ ..., N_e, \qquad j = 1, \ ..., N_p. \tag{13}$$

Under the condition that the residuals are normally distributed, the solution to Equation (12) is given by

$$\mathbf{S} = (\boldsymbol{\Psi}^\mathsf{T} \boldsymbol{\Psi})^{-1} \cdot \boldsymbol{\Psi}^\mathsf{T} \cdot \mathbf{y}, \tag{14}$$

where $\mathbf{y} = [y_1, \ ..., y_{N_e}]^\mathsf{T}$ is a vector with the outcomes of the function realizations obtained from the design experiment, $y_i = g(\boldsymbol{\xi}^{(i)})$, $i = 1 \ ... N_e$.

The PCE, as most other regression-based models, suffers from the so-called curse of dimensionality, where the number of coefficients increases dramatically for high number of dimensions and higher orders. To illustrate the typical problem size, for a PCE of order 6 and with six variables, the total number of model coefficients is $N_p = 12!/(6! \cdot 6!) = 924$, and the size of the design matrix is $N_e \times 924$. If the number of variables increases to 9, the design matrix size grows to $N_e \times 5005$. The increased number of coefficients also means that the sample size N_e needs to be larger, $N_e \gg N_p$, in order to get a determined system of equations. Working with problems of this size is still feasible on a typical computer, but the model computations slow down. However, because of the orthogonal properties of the polynomial basis, it is straightforward to determine the contribution of each polynomial term to the model variance.[15] This also gives an opportunity to achieve model reduction by applying the Galerkin approach of choosing only the fraction of the terms, which account for a sufficient contribution to the model variance (eg, 95% or

99%). Additional model reduction can be achieved by employing the least absolute shrinkage and selection operator[16] (LASSO) or other type of weighted regression in Equation (12).

For additional details regarding the calibration and use of PCE models, please refer to Sudret[15] and also Murcia et al[17] and Dimitrov et al[11] for an example application to other wind energy-related problems.
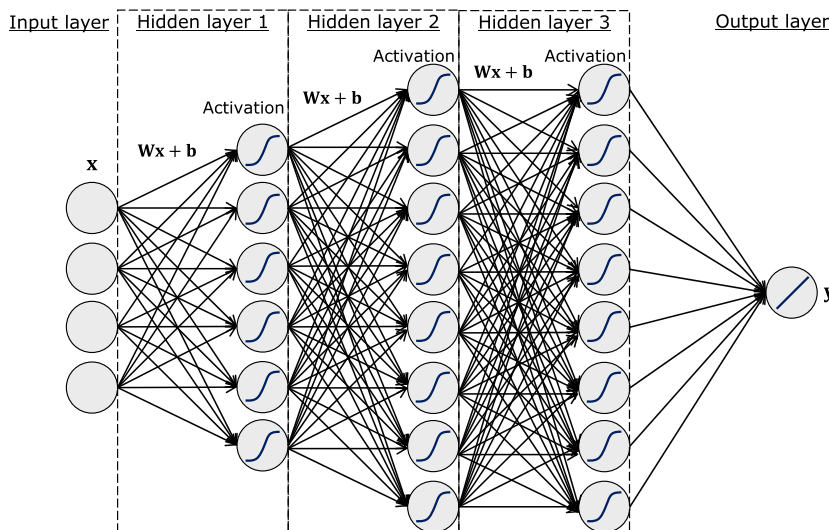
2.3.2 | Artificial neural networks

Artificial neural networks are computing models with architecture, which loosely resembles the structure and cognitive processes in animal brains. While practical applications in machine learning have been available since the 1970s,[18] ANNs gained renewed interest in the 2010s, where advances in computing power and the introduction of innovative algorithms made it possible to train complex ANN models, which offer excellent performance in a variety of tasks.[19-21] The specific architecture and the size of the ANN can vary significantly depending on the task. For the present problem where a regression model with continuous output is required, the feedforward architecture with fully connected layers (Figure 7) is adopted.

During computation, a layer in a feedforward ANN carries out two operations. First, the input vector $\mathbf{x}$ is linearly scaled to $\mathbf{z}=\mathbf{Wx}+\mathbf{b}$ by weight matrices $\mathbf{W}$ and $\mathbf{b}$. Second, the result is passed through a nonlinear activation function $a(\mathbf{z})$, and the output of the activation function serves as input to the next layer, $\mathbf{x}^{[i+1]}=a(\mathbf{z}^{[i]})=a(\mathbf{W}^{[i]}\mathbf{x}^{[i]}+\mathbf{b}^{[i]})$. Normally, a feedforward ANN will have an input layer (ie, the input variables), an output layer (the dependent variable), and one or more hidden layers between the input and the output. The choice of activation function in the output layer depends on the type of the prediction task. For obtaining a continuous regression model, the output layer is linear; ie, there is no activation. According to the universal approximation theorem,[22] a model of this architecture can approximate any real-valued function using a finite number of neurons in a single hidden layer. In the early ANN implementations, sigmoid functions (logistic and hyperbolic tangent) are the standard choice of activation functions. This has changed considerably with the employment of deep neural networks for tasks such as speech and image recognition, where the rectified linear units (ReLUs) show superior performance[19,20] and have subsequently become state of the art. However, when applied to regression modelling, ReLUs result in a piecewise linear predictive function, which is not desirable for implementations where a continuous function would be optimal. Therefore, hyperbolic tangent (tanh) is used as activation function in the present study. Various network sizes and architectures were tested, and the setup shown in Figure 7, with three hidden layers, was found to be most suitable. The number of neurons in the three hidden layers is 12, 24, and 24 respectively.

Training the ANN amounts to determining the values of the linear scaling parameters $\mathbf{W}$ and $\mathbf{b}$, which minimize the difference between the observed values of the dependent variable, $\mathbf{y}$, and the model predictions, $g(\mathbf{W},\mathbf{b},\mathbf{x})$. In the present study, the squared sum of the residuals is used as an objective function:

$$J(\mathbf{W}, \mathbf{b}) = \sum_{i=1}^{N_e} (y_i - g(\mathbf{W}, \mathbf{b}, \mathbf{x}_i))^2. \tag{15}$$

The minimization is carried out using a gradient descent approach with the Adam algorithm[23] running for 2000 epochs with learning rate 0.1, regularization 0.1, and with a batch size equal to the number of training samples.

FIGURE 7 Architecture of an example feedforward artificial neural network with three hidden layers and a linear output [Colour figure can be viewed at wileyonlinelibrary.com]

2.4 | Analytical derivatives

For applications requiring function gradients such as gradient-based optimization, an analytical expression for the partial derivatives can speed up calculations significantly as it eliminates the need for multiple function calls (as required if the derivatives are to be determined using finite-difference approach). Given the structured and relatively simple model representation in terms of a PCE and the properties of the orthogonal polynomials, it is straightforward to define analytical derivatives with respect to the standardized input variables.[24] Orthogonal polynomial families typically satisfy recursive relations for their coefficients and derivatives; this is also true for the univariate Legendre polynomials, where the derivative of a polynomial of order α satisfies the relation

$$\frac{\xi^2 - 1}{\alpha}\frac{dP_\alpha(\xi)}{d\xi} = \xi P_\alpha(\xi) - P_{\alpha-1}(\xi). \tag{16}$$

Thus, the derivative of a given multivariate polynomial Ψ_α with respect to variable ξ_i is given by

$$\frac{d\Psi_\alpha}{d\xi_i} = \left(\frac{\alpha_i}{\xi_i^2 - 1}\right)(\xi_i P_{\alpha_i}(\xi_i) - P_{\alpha_i-1}(\xi_i))\prod_{\substack{k=1 \\ k \neq i}}^{M} P_{\alpha_k}(\xi_k). \tag{17}$$

For a given expansion $\tilde{S}(\xi)$ as given by Equation (7), let us denote by $J_i \subset j = [1,...,N_p]$ as the subset of indexes for which the multivariate polynomials are dependent on variable i; ie, $\alpha_i > 0$ for all $j \in J_i$. Then, derivative of the expansion with respect to ξ_i equals

$$\frac{d\tilde{S}}{d\xi_i} = \sum_{J_i} S_j \frac{d\Psi_{\alpha,j}}{d\xi_i} = \sum_{J_i} S_j \left(\frac{\alpha_{i,j}}{\xi_i^2 - 1}\right)(\xi_i P_{\alpha_{i,j}}(\xi_i) - P_{\alpha_{i,j}-1}(\xi_i))\prod_{\substack{k=1 \\ k \neq i}}^{M} P_{\alpha_{k,j}}(\xi_k). \tag{18}$$

Computing the derivatives of the surrogate model using Equation (18) is straightforward and often significantly faster than using finite differences. This can be further facilitated by storing all the polynomial orders prescribed by Equation (17) in a 2-D matrix of size (M,N_p), which is invariant to the actual values of ξ and thus only needs to be computed once. In this case, a derivative computation only requires evaluation of Legendre polynomials of the prescribed orders, followed by simple summation and product operations.

A similarly effective gradient computation can be established for ANNs. As the ANN model represents a sequence of nested simple functions with known derivatives, gradients can be computed using the chain differentiation rule, which for a function $y = g(z)$ where $z = f(x)$ states that

$$\frac{dy}{dx} = \frac{dy}{dz}\cdot\frac{dz}{dx}. \tag{19}$$

The above is routinely used in training of ANN models, where the so-called back-propagation approach is used to compute the gradient of the cost function with respect to model parameters. Back propagation can also be applied in order to determine gradients of the dependent variable with respect to the inputs. Noting that the output of layer i serves as the input for layer $i+1$, one back-propagation step is defined as

$$\frac{dy}{d\mathbf{x}^{[i]}} = \mathbf{W}^{[i],T}\left(\frac{dg(\mathbf{z}^{[i]})}{d\mathbf{z}}\cdot\frac{dy}{d\mathbf{x}^{[i+1]}}\right). \tag{20}$$

In the above, T denotes matrix transpose, the (.) sign denotes element-wise multiplication, and $dg(\mathbf{z}^{[i]})$ is the derivative of the activation function of the ith layer with respect to $\mathbf{z}^{[i]} = \mathbf{W}^{[i]}\mathbf{x}^{[i]} + \mathbf{b}^{[i]}$. For the hyperbolic tangent used as hidden layer activation function in the present study, $dg(\mathbf{z}^{[i]})/d\mathbf{z} = 1 - \tanh^2\mathbf{z}^{[i]}$. For the output layer in a regression model, the activation function is $g(\mathbf{z}) = \mathbf{z}$ and subsequently $dg(\mathbf{z})/d\mathbf{z} = 1$. Starting with the output layer and applying the chain rule with consecutive evaluations of Equation (20), the value of $dy/d\mathbf{x}^{[0]}$ is determined, which is the derivative of the dependent variable with respect to the input variables.

2.5 | Seed-to-seed uncertainty

Aeroelastic load simulations use the so-called turbulence boxes as inputs, which are random realizations of three-dimensional, stationary, and homogeneous turbulent wind fields. Even under the same input conditions and having the same statistical properties, the time series in the turbulence boxes will differ from realization to realization, which we denote as seed-to-seed uncertainty. This variability manifests as uncertainty in the resulting loads, and it can be as large in magnitude as the effect of significant change in input conditions such as wind shear or turbulence length scale.[25] Normally, seed-to-seed uncertainty is accounted for by running multiple or long-period load simulations under the same conditions. The total simulation period should be at least 1 hour as recommended by the IEC61400-1 standard, [26] but often, a longer period may be necessary

for obtaining reliable load estimates.[27] When building surrogate models, the seed-to-seed uncertainty can also be accounted for by running simulations with multiple random seeds at each model training point. Then, two surrogate models may be fit—one to the sample means from each training point and another to the standard deviations. The former is used for obtaining a mean load estimate at certain conditions, while the latter can be used to assess the uncertainty and build confidence intervals for the load estimates.

2.6 | Input variable space

Calibration of the surrogate models requires a training data set that is sufficiently large to allow solving the mathematical equations for determining the model coefficients. The training set should also cover the variable space to the best extent possible. For the present problem, a stochastic collocation approach is adopted similarly to Dimitrov et al and Murcia et al.[11,17] The sample space is represented in terms of independent, uniformly distributed variables in the range [0,1], which represent the cumulative distribution functions of the physical variables. The mapping from physical to uniform space is done by means of a Rosenblatt transformation.[28] The Rosenblatt transformation allows taking dependence into account by representing the joint distribution of the physical variables as a series of conditional approximations. The first variable is considered independent, and subsequent variables can be conditionally dependent on preceding variables but not on variables lower in the list. As typical for wind energy-related applications, the wind speed can be considered as the independent variable, as most other variables are conditionally dependent on it. For the present problem, the turbulence σ_u, the wind shear exponent α, the significant wave height H_s, and the wave peak period T_p are considered dependent on the wind speed. In addition, α is considered conditionally dependent on σ_u, and T_p is conditional on H_s. Conditional independence is assumed for all other variables (eg, it is assumed that the water depth is not dependent on wind shear). The conditional distributions and variable ranges for wind speed, turbulence, and wind shear used in the present study are based on Dimitrov et al.[11] A suggested joint distribution model for wind speed, significant wave height, and wave peak period is the one provided by Johannessen et al.[29] An example with bounds for all presently considered variables is given in Table 2. Having the joint distribution and variable bounds in place, a training sample can be generated using a pseudo-Monte Carlo (pseudo-MC) approach based on a low-discrepancy sequence (eg, Halton or Sobol sequence). The low-discrepancy series offer more uniform coverage of the sample space and faster convergence of the sample statistics.[30] Since these sequences are deterministic and nonrepetitive, they offer easy means of experimental design enrichment (eg, for increasing accuracy) by simply adding extra points to the sequence when necessary. Figure 8 shows a pseudo-MC sample of size 1024 for three variables: wind speed, turbulence, and wind shear, including the variable bounds as defined by Table 2. Since the wind speed is independent of other variables, the wind speed sample simply represents the marginal distribution (a truncated Weibull), and in Figure 8, it is represented by its histogram.

2.7 | Algorithms for model training and for determining wake variables

In a typical application scenario of the presently discussed approach, a large number of load simulations are carried out in order to train a surrogate model for a particular wind turbine type. The load simulations include conditions covering the entire input variable space in a representative way,

TABLE 2 Bounds of variation of the variables considered

Variable	Lower Bounds	Upper Bounds
$\bar{u}$	$\bar{u} \geq 4$ m/s	$\bar{u} \leq 25$ m/s
σ_u	$\sigma_u \geq 0.025 \cdot \bar{u}$ (m/s)	$\sigma_u \leq 0.18\left(6.8 + 0.75U + 3\left(\frac{10}{\bar{u}}\right)^2\right)$ (m/s)
α	$\alpha \geq \alpha_{ref,LB} - 0.23\left(\frac{\bar{u}_{max}}{\bar{u}}\right)\left(1 - \left(0.4\log\frac{R}{z}\right)^2\right)$	$\alpha \leq \alpha_{ref,UB} + 0.4\left(\frac{R}{z}\right)\left(\frac{\bar{u}_{max}}{\bar{u}}\right)$
H_s	Deterministic	Deterministic
T_p	Deterministic	Deterministic
Δ	$\Delta \geq -20°$	$\Delta \leq 20°$
Z_w	$Z_w \geq 6$ m	$Z_w \leq 14$ m
R_D	$R_D \geq 6.5D$	$R_D \leq 25D$
θ	$\theta \geq -20°$	$\theta \leq 20°$

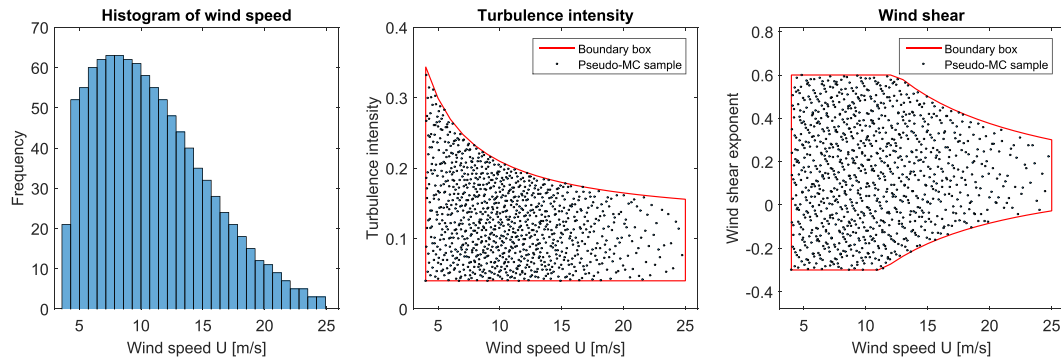Note. All environmental variables are defined as statistics over 10-minute reference period.

Where

[a]R is the rotor radius, D the rotor diameter;

[b]$\alpha_{ref,LB}$=0.15, $\alpha_{ref,UB}$=0.22 are reference wind shear exponents at 15m/s wind speed;

[c]$\bar{u}_{max} = 25$ m/s is the upper bound of the wind speed;

[d]ϕ is the reference latitude (here chosen as 50°).

FIGURE 8 Illustration of the variable space with bounds for wind speed, turbulence, and wind shear [Colour figure can be viewed at wileyonlinelibrary.com]

as shown in Section 2.6. When deployed, the surrogate model can provide almost instantaneous load and power maps for various wind farm layouts and wind conditions as long as the same wind turbine type is used. The process of building and training the surrogate model is similar to the one defined in Dimitrov et al[11] and is outlined below in terms of the necessary steps:

1. Define the variable space: which variables are to be included in the training, what are their ranges, and if there are any dependencies between variables (eg, the turbulence range may depend on the wind speed).

2. Generate a random sample covering the input variable space in a uniform way (see Section 2.6).

3. Carry out load simulations for each point in the sample set defined in point 2. In order to account for seed-to-seed uncertainty (Section 2.5), carry out multiple simulations at each sample point and collect the statistics (mean and standard deviation) of the output variables for each point.

4. Train a surrogate model that based on the inputs defined in point 2 can predict the outcomes of the simulations from point 3.

Having designed and trained a surrogate model for the wake-induced load effects, producing a load map over the wind farm requires defining the wake-related variable inputs based on the farm geometry for each turbine of interest. Below, the step-by-step procedure is described, taking as an example the geometry of the Horns Rev 1 (HR1) offshore wind farm in the Danish North Sea (Figure 9) and looking at turbine T55 in the park in particular.

1. Choose a turbine for analysis (let us denote it by T_i). Compute relative distances and azimuth angles between the position of T_i and all other turbine positions in the farm.
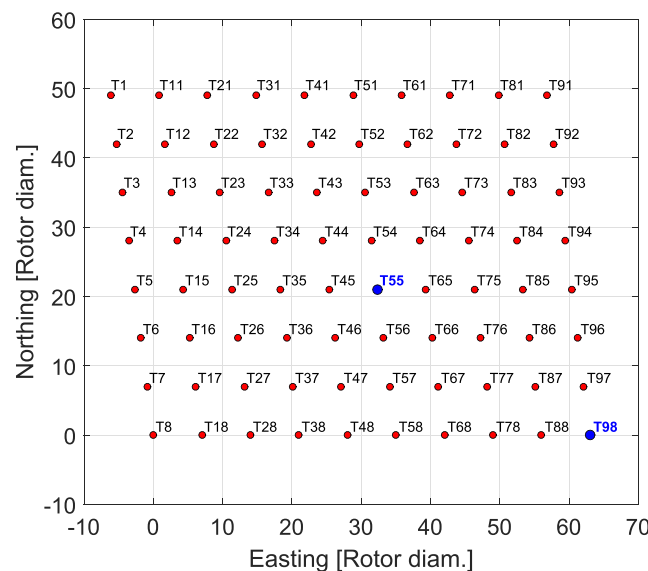

FIGURE 9 Map of the HR1 wind farm with turbine names and spacing in rotor diameters. Turbine positions used in the analysis (T55 and T98) are highlighted [Colour figure can be viewed at wileyonlinelibrary.com]

2. Group turbines in rows by relative azimuth–ie, all turbines for which the relative azimuth is within a small range (eg, ±0.5°), are considered as members of the same disturbing row. Turbines that do not become included in rows are considered as single (one row) wake sources. The result is a set of possible wake disturbances with known azimuth and distance to the first turbine. As shown in Section 2.2, wake effects caused by any given wake disturbance can be experienced for alignment angles of up to ±20°. This means that for some angles, there is potential overlap of possible wake sources, as shown by the raw "disturbance rose" plotted on the left hand side of Figure 10.

3. Split the azimuth circle in small (eg, 2° step) bins. For each azimuth bin, evaluate the probability of observing wake interaction from all possible disturbances identified in step 2, by using Equation (6). For each azimuth bin, assign the wake-related variables $[R_D, \theta, N_{rows}]$, which correspond to the wake disturbance with highest probability. The result is a disturbance rose without overlaps, as shown on the right-hand plot in Figure 10. The resulting values of $[R_D, \theta, N_{rows}]$ over the full azimuth circle are shown in Figure 11.

4. Use the load surrogate model to compute mean (and if available, standard deviation) of loads as function of $[R_D, \theta, N_{rows}]$, as well as any other relevant variables from the input space (Equation 1).

If the load surrogate model is sufficiently fast, a probability-weighted sum of the loads from all possible disturbances could be employed. Alternatively, the computation of wake source probabilities can be avoided by merging points 3 and 4 in the above procedure and estimating wake-induced loads for all possible disturbances, taking a relative angle range of $\approx \pm 20°$ for each disturbance. These are shown as thin colored lines in Figure 12. Then, for each azimuth bin, the wake variable selection corresponds to the disturbance source for which the surrogate model predicts the highest loads. This results in the thick black line shown in Figure 12. As it was shown earlier (Figure 3), the wake-induced load excess follows the wake incidence probability fairly well, which justifies the procedure; this approach is also conservative since in case of multiple disturbances, the one causing the highest loads is selected. In the case study presented in the following section, the latter approach with selecting the disturbance corresponding to the highest loads has been used.

3 | DEMONSTRATION

In the following, an example application of the wake interaction mapping procedure is shown based on the HR1 wind farm. A surrogate model for wake-induced loads is trained on a generic data set and used to compute loads experienced by turbines 55 and 98, which are then compared with loads computed by a full wind farm simulation using the DWM model[3] and Hawc2.[31]
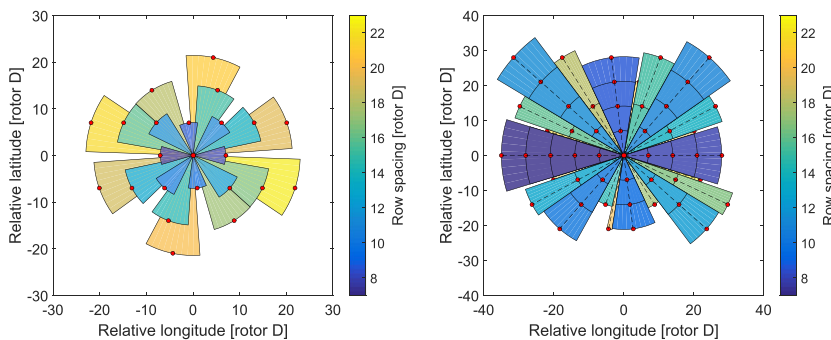
FIGURE 10 Left: disturbance rose for turbine T55 in the HR1 wind farm—showing full range of possible angles with disturbance from any turbine within 23 rotor diameters distance. Right: complete disturbance rose for T55 without overlaps and with multiple rows of disturbing turbines. The color coding denotes distance between T55 and disturbing turbines [Colour figure can be viewed at wileyonlinelibrary.com]
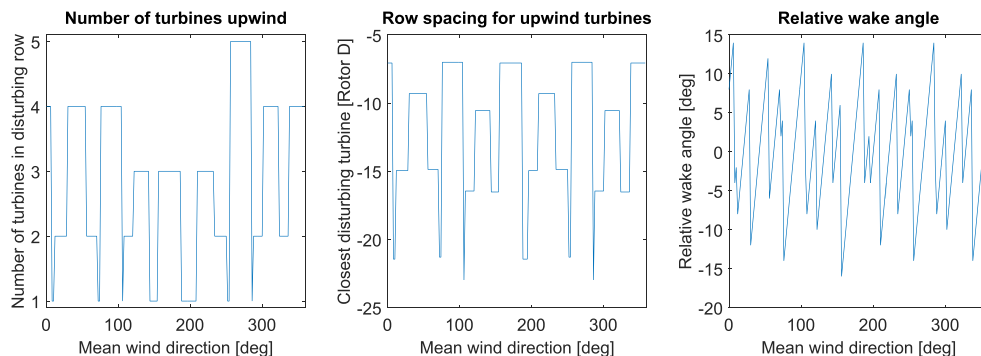
FIGURE 11 Illustration of the wake disturbance conditions as function of wind direction for turbine T55 [Colour figure can be viewed at wileyonlinelibrary.com]
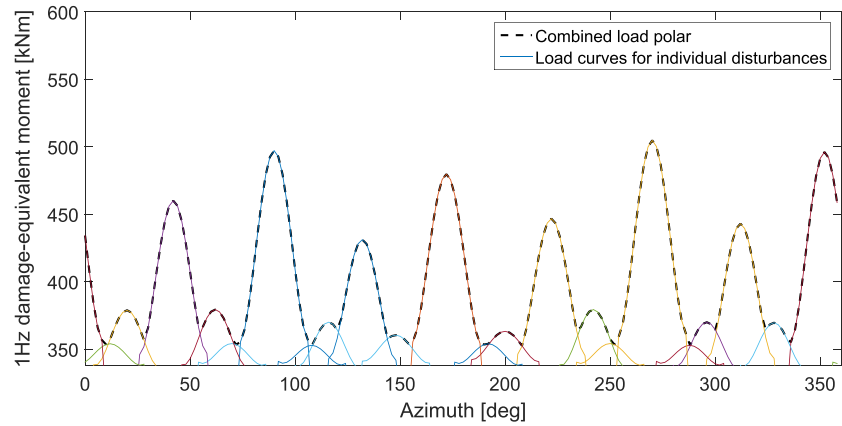
FIGURE 12 Load polar for blade root fatigue equivalent moment for turbine T55 at HR1 for mean wind speed of 13 m/s [Colour figure can be viewed at wileyonlinelibrary. com]

3.1 | Load simulation setup

In order to produce a data set for training a load surrogate model, a large database with aeroelastic load simulations is generated. The database covers the variable space described in Section 2.6, and the sample generation is done using a pseudo-MC approach. The variable bounds used for generating the pseudo-MC samples are given in Table 2. For simplicity, two of the variables (the significant wave height and the wave peak period) are taken as deterministic and equal to their mean values conditional on the mean wind speed as in the model prescribed by Johannessen et al.[29] This reduces the dimension of the required surrogate model by 2, which also lowers the number of simulations required for proper calibration. Notably, the variable N_{rows} describing the number of disturbing turbines is not included in Table 2. This is because the dependence between the number of disturbing turbines and the loads is strongly nonlinear—especially a change from zero rows (ie, no disturbance) to one disturbing turbine is very abrupt. Such a behaviour is not well replicated by models based on polynomial theory. To overcome this problem, separate PCEs are trained, where each expansion is based on a data set with just one value of N_{rows}. While it is not strictly necessary to use this approach with ANNs, for consistency, the same procedure as with the PCE is used. A 5000-point pseudo-MC sample over all other variables (as shown in Table 2) is generated for each value of N_{rows} considered. Thus, the final surrogate model consists of several submodels covering the range of discrete values of N_{rows}. In the present case, turbine T55 can experience at most five disturbing turbines aligned in a row; thus, six values of N_{rows} are required, from 0 to 5. Six simulations with different turbulence seeds are carried out for each sample point; therefore, a total of 180 000 10-minute aeroelastic simulations are carried out in order to generate the entire training data set.

Load simulations are carried out using a Hawc2 aeroelastic model of a generic 2-MW wind turbine (same power rating as the V80 wind turbines installed at HR1), running with the DTU basic wind turbine controller.[32] The wake-induced effects including the meandering are introduced by generating three turbulence boxes for each simulation: a large-scale turbulence field used to establish the meandering paths of the wake deficits, a standard turbulence box representing the ambient wind field over the rotor area, and a microturbulence box representing the wake deficit, which is superimposed on the ambient turbulence box. The relative position of the microturbulence box within the ambient turbulence box is determined from the wake meandering path computed from the large-scale turbulence box. Six turbulence seeds are used per MC sample, in order to reduce the seed-to-seed uncertainty, although, as discussed in Dimitrov et al,[11] using a large MC sample has a similar effect on the uncertainty in the surrogate model estimations. DELs are computed for each load simulation, and the final load value corresponding to a given sample point is taken as the mean DEL from all valid simulations with the same conditions. Twelve load channels are considered: fore-aft and side-side bending moments at seabed, mean sea level, tower base, and tower top; yaw moment; shaft torsion; and blade root flapwise and edgewise bending moments. Due to no soil properties information available, an apparent fixity model is used for the turbine monopiles. The monopile is fixed rigidly to a point approximately 10 m below the seabed, and the stiffness properties of the monopile section between the fixed point and the seabed are tuned to match the expected first natural frequency of the turbine system. Variation in depth is then achieved by modifying the length of the piece of monopile between the seabed and the mean sea level.

3.2 | Effect of individual input variables

A quick surrogate model offers a convenient possibility of exploring the dependence of the turbine loads on different input variables. Figure 13 shows the surrogate model predictions for different component DEL as function of water depth, for mean wind speed of $\bar{u} = 10$ m/s, turbulence intensity of 0.1, wind shear exponent α=0.1, and no wake disturbance (N_{rows}=0). The results show a linear dependence between water depth and loads for the parts of the structure below sea level, which can be attributed to the increase of moment arm of the rotor thrust force when the
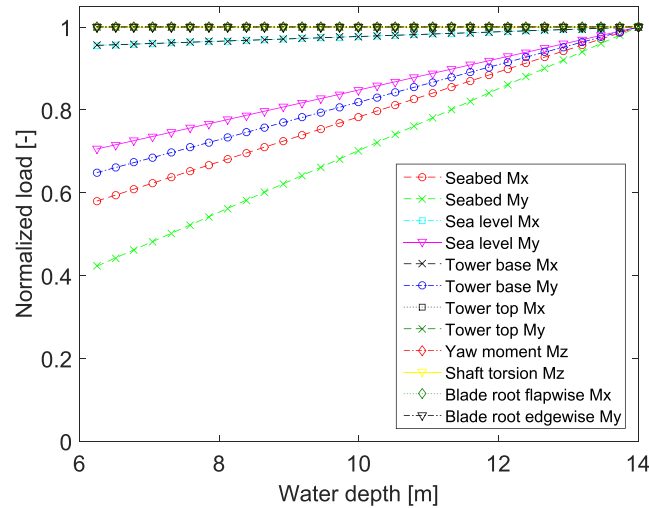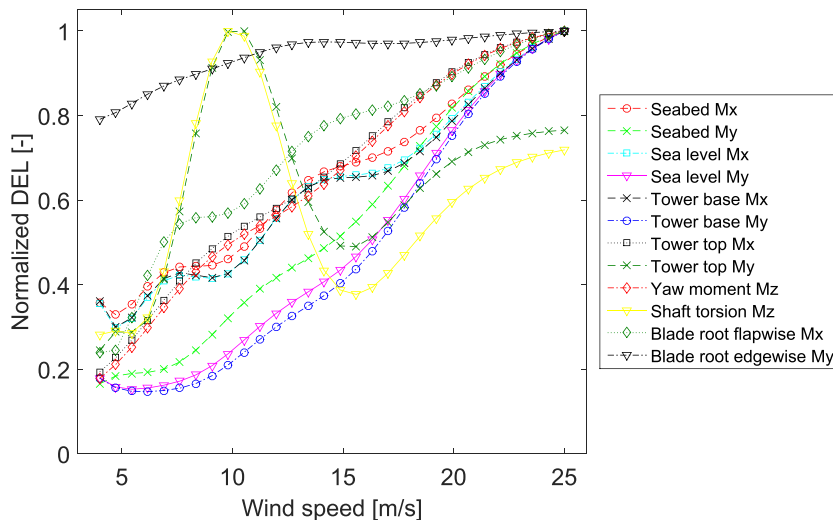
FIGURE 13 Surrogate model-based prediction of damage-equivalent loads as function of water depth, at 10m/s mean wind speed, for different wind turbine components [Colour figure can be viewed at wileyonlinelibrary.com]

depth increases. This presentation of expected behaviour reinforces the confidence that the surrogate model can capture the important aspects of the wind turbine behaviour. Figure 14 shows the dependence of 12 DEL components on wind speed as predicted by a PCE, with turbulence equal to the 50% quantile as defined by the joint distribution of $\bar{u}$ and σ_u from Table 2, and with no wake effects. While for the majority of the loads, there is a monotonic increase of load with wind speed, the shaft torsion and some thrust-related loads (eg, the blade root flapwise moment) show quite nonlinear behaviour. While this feature of the surrogate model may seem unusual, it simply mimics the physical behaviour of the loads. This can be observed in Figure 15, where the surrogate model predictions for shaft torsion and blade root flapwise moments are compared with points from the model training sample. Despite the large nonlinearity in the response, the surrogate model provides a good prediction, again showing that the surrogate model should be able to provide a valid representation of the turbine system. With regard to the wake-effect modelling capabilities of the surrogate model approach, Figure 5 demonstrates the load predictions for different numbers of disturbing turbines arranged in a row. As discussed, the bell-shaped load curve looks similar to the probability distribution of the wake position as defined by the DWM model, and the behaviour is also similar to that of the equivalent turbulence as defined by the Sten Frandsen model.[1] The loads shown in Figure 5 increase with the number of turbines in a row, which is expected for wind speeds above rated, where subsequent wake deficits are superimposed by the DWM model implemented in Hawc2.

3.3 | Fatigue load estimation results

To assess the performance of the load surrogate in predicting the behaviour of the DWM model, reference simulations for turbines T55 and T98 with the DWM model and including all other turbines in the HR1 wind farm as wake sources are carried out. For obtaining a good resolution and

FIGURE 14 Surrogate model-based prediction of damage-equivalent loads as function of wind speed for different wind turbine components [Colour figure can be viewed at wileyonlinelibrary.com]
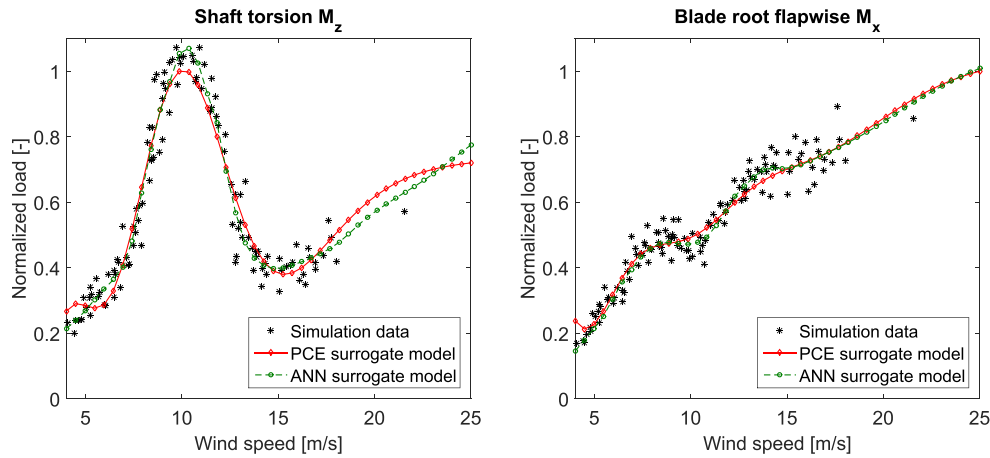
FIGURE 15 Surrogate model-based prediction of damage-equivalent loads as function of wind speed, compared to points from training sample. Left: main shaft torsion; right: blade root flapwise bending moment [Colour figure can be viewed at wileyonlinelibrary.com]

accuracy in the reference case, wind directions from 0° to 360° are simulated in 2° steps, and at each direction, 12 independent turbulent realizations are used. The simulations are repeated for wind speeds from 5 to 25 m/s in 2-m/s steps, the average turbulence intensity varies from 0.07 to 0.08 for different wind speeds corresponding to the site-specific conditions, and the wind shear is taken as 0.1. The wave conditions are represented by the conditional mean T_p and H_s, and the water depth is taken as 10.2 m for T55 and 8.2 m for T98. For each azimuth bin, the mean DEL from all realizations is evaluated and compared with the mean prediction from the surrogate model evaluated including the wake-related variable values shown in Figure 11. The blade root flapwise DEL prediction comparison for T55 at 5-, 9-, 13-, and 17-m/s wind speed is shown in Figure 16. Figure 17 shows the same comparison for tower base fore-aft loads at T98.

A quantitative comparison is also carried out by computing the normalized root mean square error (NRMSE) between the actual load values y and the surrogate model predictions $\tilde{y}$:

$$NRMSE = \frac{1}{\bar{y}}\sqrt{\frac{\sum_{i=1}^{N_e}(\tilde{y}_i-y_i)^2}{N_e}}, \tag{21}$$

where $\bar{y}$ is the mean of the observations of the dependent variable and N_e is the number of observations. Table 3 lists the combined NRMSE computed over all wind directions and for wind speeds from 5 to 21 m/s, for eight load channels and for both the PCE and ANN surrogate model approaches. In order to assess the seed-to-seed uncertainty, the standard error of the mean predictions from multiple realizations is estimated for the full-farm DWM model as $SE = \sigma_y/\sqrt{M}$, where M=12 is the number of independent turbulent realizations for each simulated condition and σ_y is the standard deviation of the resulting fatigue DEL. To allow adequate comparison with the overall error described by the NRMSE, the SE is normalized by the mean of the observed load, $\bar{y}$, and the results are also listed in Table 3.

As one of the potential applications of the present approach is for wind farm load mapping, the performance in estimating lifetime fatigue loads is also assessed. The lifetime damage-equivalent loads are estimated with a probability-weighted sum of the individual model predictions, assuming uniform wind rose and Weibull-distributed wind speed with distribution parameters A=11.28 and k=2 (ie, IEC class 1 conditions). The relative difference between the surrogate model predictions and the full DWM model simulations is listed in Table 3.

The presently used surrogate models have the capability to estimate load in both free wind and wake conditions, which can be used to determine the relative contribution of the wake-induced effects to the total accumulated damage over the turbine lifetime. This is listed in Table 3 as the ratio between the lifetime DEL computed with wake effects to the lifetime DEL computed by assuming a wind turbine operating in free wind 100% of the time. The load components that are in general more correlated with the longitudinal wind speed and the thrust force (such as the blade root flapwise moment and the tower base fore-aft bending moment) are significantly affected by the wake-induced disturbance, while for blade edgewise moments, the effect of wakes is much smaller.

4 | DISCUSSION

The load prediction results showed up to 3% (ANN) and 10% (PCE) difference in the estimated long-term DEL loads between the surrogate model approach and the full-farm DWM simulation. This difference is due to a combination of realization-to-realization (seed-to-seed) uncertainty in the full-farm simulations and bias due to possible inaccuracy in the surrogate model representation. As seen from Table 3, the seed-to-seed
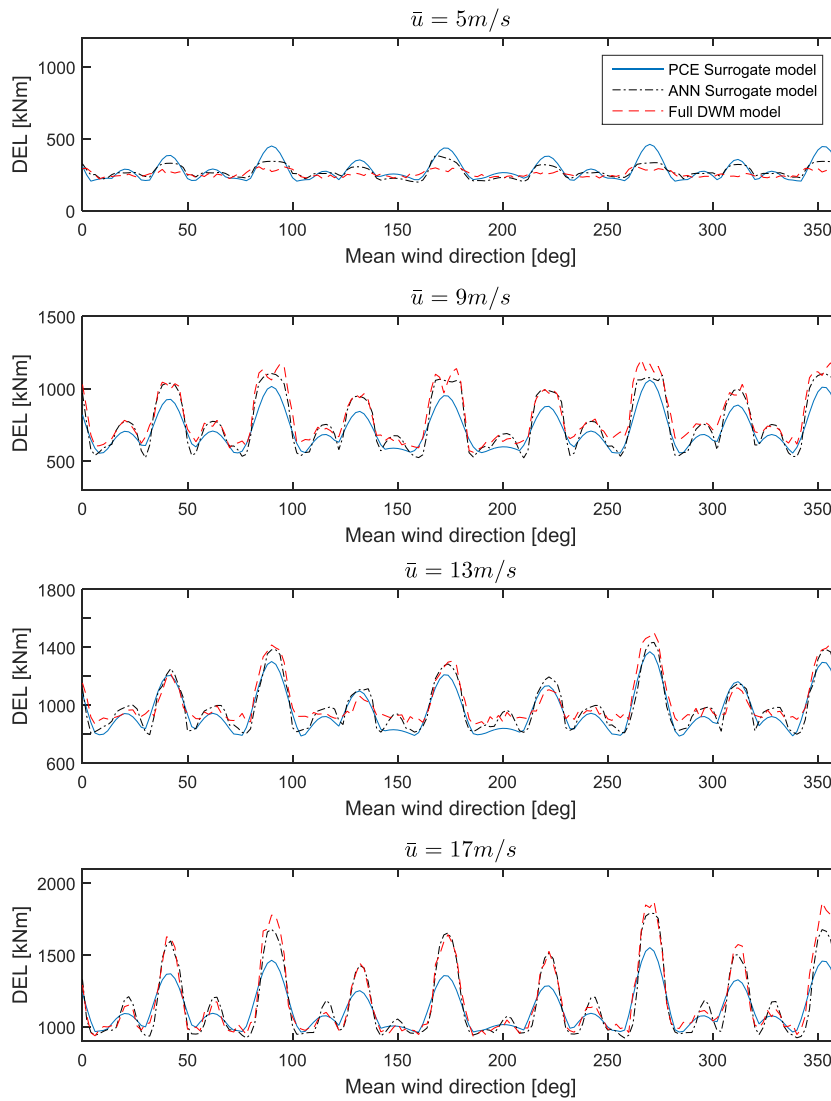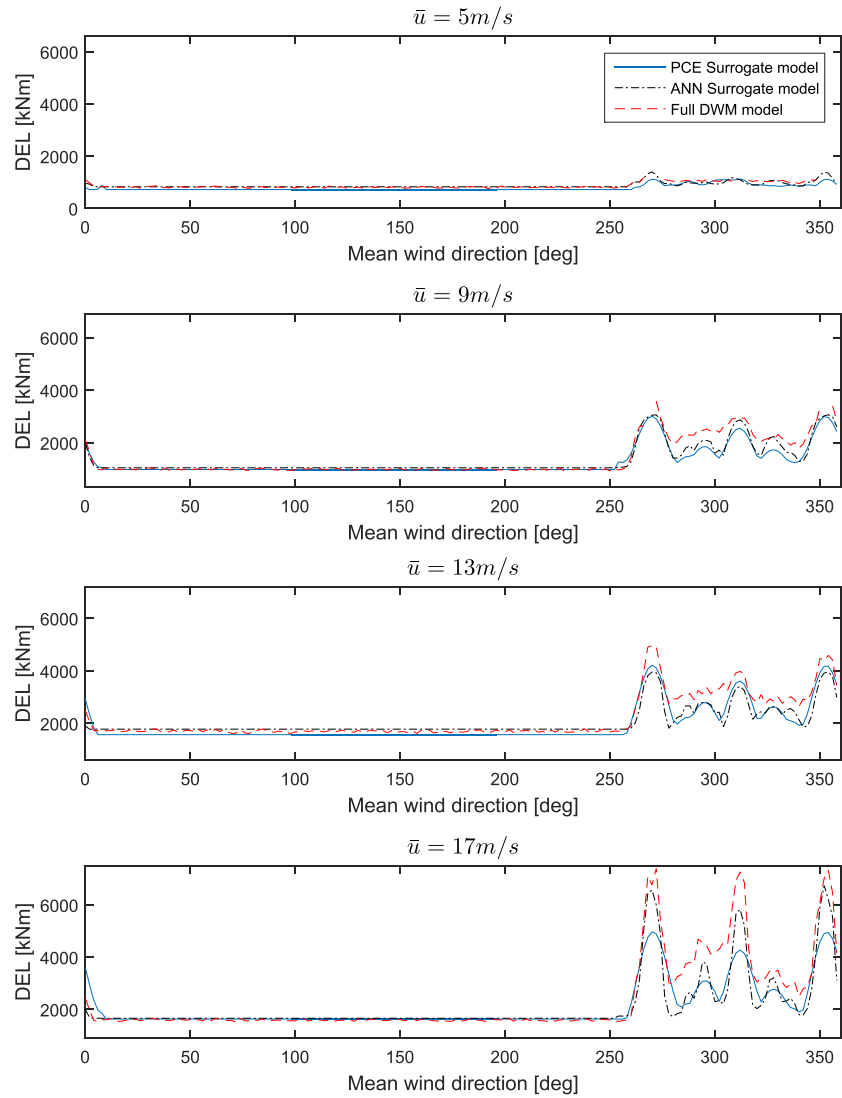
FIGURE 16 Comparison of the wake load surrogate model performance against full-farm dynamic wake meandering model computations: prediction of blade root flapwise fatigue loads for T55 at HR1 wind farm [Colour figure can be viewed at wileyonlinelibrary.com]

uncertainty accounts for about 10% to 30% of the total error. This uncertainty can be at least partly eliminated by increasing the number of seeds per realization for the full-farm DWM simulation. Nevertheless, for the purposes of, eg, optimization, it is most important that the surrogate model properly predicts the load differences caused by changes in external conditions. The distribution of the absolute error of the surrogate models is relatively uniform over the range of input variables. However, as the loads experienced by the turbines increase significantly with the wind speed, the relative error at low wind speeds is much larger than at high wind speeds—NRMSE of up to 30% of the mean value was observed for some load channels (eg, shaft torsion) at $\bar{u} = 5$ m/s, while for $\bar{u} \geq 21$ m/s, the NRMSE is in the order of 1% and does not exceed 5% of the mean value. This means that in case of using the surrogate model in an optimization problem, the confidence in the solutions at low wind speeds will be smaller than at high wind speeds. This may be partially alleviated by increasing the amount of training data.

A comparison of Figures 16 and 17 shows that with both the PCE and the ANN surrogate models, the load peaks due to wake effects are captured better for T55 than for T98. The surrogate models have been trained on simulations with up to five rows of disturbing turbines upwind, while T98 can experience up to nine turbines upwind in certain wind directions. These results indicate that, especially at high wind speeds, the dependency between wake-induced loads and number of upwind turbines does not level off completely even when considering five rows of turbines upwind.

The nonnegligible seed-to-seed uncertainty means that the output of high-fidelity load simulation runs will in general be a noncontinuous function of the inputs. The surrogate model approach provides a continuous representation of the load response; however, for some applications, the seed-to-seed uncertainty needs to be accounted for, as it arises from a real physical phenomenon—the turbulent inflow. As stated in Section 2.5, the seed-to-seed uncertainty may be mapped by means of training an additional surrogate model for predicting the standard deviation of load outputs. It should be noted that this approach is sufficient for uncertainty analysis relying on first- and second-order statistics such as confidence intervals but most likely not sufficient for estimating small probabilities for rare events (for example, for structural reliability analysis). The latter would require better accuracy with significantly larger amount of data to properly characterize higher-order statistical moments.

FIGURE 17 Comparison of the wake load surrogate model performance against full-farm dynamic wake meandering model computations: prediction of tower base fore-aft fatigue loads for T98 at HR1 wind farm [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Errors between DEL predictions for wake-mapping surrogate model and full wind farm dynamic wake meandering (DWM) simulations

Criterion	$M_{x,TB}$	$M_{x,TT}$	$M_{y,TT}$	$M_{z,TT}$	$M_{z,S}$	$M_{x,B}$	$M_{y,B}$
Seed-to-seed uncertainty NRMS	0.045	0.025	0.040	0.033	0.051	0.029	0.004
NRMSE, PCE	0.158	0.086	0.165	0.082	0.181	0.135	0.021
NRMSE, ANN	0.124	0.084	0.138	0.106	0.124	0.084	0.019
Lifetime DEL prediction error, PCE	0.060	−0.009	0.072	0.031	0.093	0.095	0.005
Lifetime DEL prediction error, ANN	0.008	−0.006	−0.004	−0.029	0.014	0.005	−0.010
Wake-induced lifetime DEL increase	1.742	1.095	1.839	1.191	1.988	1.387	1.029

Note. Load abbreviations correspond to the following: $M_{x,TB}$, tower base fore-aft loads; $M_{x,TT}$, tower top tilt moment; $M_{y,TT}$, tower top side-side moment; $M_{z,TT}$, tower top yaw moment; $M_{z,S}$, shaft torsion; $M_{x,B}$, blade root flapwise moment; and $M_{y,B}$, blade root edgewise moment.

Abbreviations: ANN, artificial neural network; DEL, damage-equivalent fatigue load; NRMSE, normalized root mean square error; PCE, polynomial chaos expansion.

An important factor influencing the adequateness of the surrogate models is the layout of the wind farm and the variables that are used to parameterize the wake-induced effects. The present choice of variables is most suitable for farms with regular layouts where it can be expected that upwind turbines form well-aligned rows. The same variable set can also be useful for farms with fully random layout if the separation between turbines is large enough so that a single predominant source of disturbance can be identified for any wind direction. For other types of layouts, the same overall concept could be used, but most likely, additional variables would be needed to achieve sufficiently adequate representation. Such variables could, for example, include wake sources in more than one direction upwind in order to take the effect of wake mixing into account.

In the present study, the 10-minute average wind speed, turbulence, and wind shear are chosen as the variables that describe the ambient wind conditions. There may be other variable choices that are equally or more suitable for the purpose—eg, the atmospheric stability that has a complex relationship with wind speed, turbulence, and wind shear.[33] While combinations of wind shear and turbulence usually correspond to certain stability conditions and can be used as a proxy, it may be more convenient to directly include stability as a variable as demonstrated in Keck et al[34] and Larsen et al.[35]

The present paper considers methods for load mapping using the DWM model. The same setup and models can in principle be used for mapping power production; however, power prediction requires very narrow confidence bounds, which may be difficult to achieve with the DWM without increasing the computation efforts. On the other hand, similar approach based on surrogate models can be used for mapping of power production when combined with wind farm flow models as the ones seen in, eg, Murcia Leon et al.[36]

The DWM model performance has been validated against numerous measured load data sets, while in the present study, a surrogate model–based approach is validated against the DWM model predictions. However, direct validation of the performance of the surrogate model against measured data was not currently possible. This is an obvious potential topic for future research.

Two approaches for calibrating efficient surrogate models using machine learning were presented: PCE and ANN. The ANN outperformed PCE in terms of prediction accuracy and computational speed; however, ANNs are more prone to overfitting and seem to require more data for achieving optimum performance. The ANN model may be the preferred choice in most contexts unless the training data set is not sufficiently large to eliminate overfitting. Both the PCE and ANN provide possibility for obtaining analytical derivatives, which is a useful feature for optimization applications and sensitivity analysis.

5 | CONCLUSIONS

A procedure for mapping wake-induced load predictions computed with the DWM model to a computationally efficient surrogate model approximation is defined and demonstrated. The following was found:

- The wake incidence probability as prescribed by the DWM model was computed, and the load behaviour was found to coincide well with the probability density curve.

- An algorithm for determining the values of the descriptive variables for wake effect based on the farm layout was defined.

- Surrogate models based on PCEs and ANNs and including three wake-related variables: wind-wake alignment angle, row spacing, and number of disturbing turbines, are able to predict the long-term effect of wake-induced loads with less than 10% error while giving a good indication of the change of loads as function of environmental conditions.

- The analytical derivatives of the PCE and feedforward ANN models were defined.

- The resulting load assessment procedure provides continuous, differentiable output with known analytical derivatives and can be used for applications such as wind turbine layout optimization, estimation of turbine lifetime, and uncertainty analysis.

ORCID

Nikolay Dimitrov https://orcid.org/0000-0003-1325-4512

REFERENCES

1. Frandsen ST. Turbulence and turbulence-generated structural loading in wind turbine clusters. Riso-R-1188, Risoe National Laboratory; 2007.

2. Keck R-E, de Maré M, Churchfield MJ, Lee S, Larsen G, Madsen HA. Two improvements to the dynamic wake meandering model: including the effects of atmospheric shear on wake turbulence and incorporating turbulence build-up in a row of wind turbines. *Wind Energy*. 2015;18(1):111-132.

3. Larsen G, Madsen HA, Thomsen K, Larsen TJ. Wake meandering: a pragmatic approach. *Wind energy*. 2008;11(4):377-395.

4. Larsen TJ, Madsen HA, Larsen G, Hansen KS. Validation of the dynamic wake meander model for loads and power production in the Egmond aan Zee wind farm. *Wind Energy*. 2013;16(4):605-624.

5. Larsen Torben J., Larsen Gunner C, Aagaard Madsen H, Petersen SM. Wake effects above rated wind speed. An overlooked contributor to high loads in wind farms. In: Scientific Proceedings. EWEA Annual Conference and Exhibition 2015. European Wind Energy Association (EWEA); 2015:95-99.

6. Larsen TJ, Larsen G, Pedersen MM, Enevoldsen K, Madsen HA. Validation of the dynamic wake meander model with focus on tower loads. *Journal of Physics: Conference Series*. 2017;854(1):012027.

7. Reinwardt I, Gerke N, Dalhoff P, Steudel D, Moser W. Validation of wind turbine wake models with focus on the dynamic wake meandering model. *Journal of Physics: Conference Series*. 2018;1037(7):072028 (10 pp.)

8. Keck R-E, Undheim O. A pragmatic approach to wind farm simulations using the dynamic wake meandering model. *Wind Energy*. 2015;18(9):1671-1682.

9. Galinos C, Dimitrov N, Larsen TJ, Natarajan A, Hansen KS. Mapping wind farm loads and power production—a case study on Horns Rev 1. *Journal of Physics: Conference Series*. 2016;753(3):032010.

10. Mann J. The spatial structure of neutral atmospheric surface-layer turbulence. *Journal of Fluid Mechanics*. 1994;273:141-168.

11. Dimitrov N, Kelly M, Vignaroli A, Berg J. From wind to loads: wind turbine site-specific load estimation with surrogate models trained on high-fidelity load databases. *Wind Energy Science*. 2018;3(2):767-790.

12. Schröder L, Dimitrov N, Verelst DR, Sørensen JA. Wind turbine site-specific load estimation using artificial neural networks calibrated by means of high-fidelity load simulations. *Journal of Physics: Conference Series*. 2018;1037(6):062027.

13. Ghanem RG, Spanos PD. *Stochastic Finite Elements—A Spectral Approach*. Berlin: Springer; 1991.

14. Xiu D, Karniadakis GE. The Wiener-Askey polynomial chaos for stochastic differential equations. *Journal of Scientific Computing*. 2002;191(43):4927-4948.

15. Sudret B. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and Systems Safety*. 2008;93:964-979.

16. Tibshirani R. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society, Series B*. 1996;58(1):267-288.

17. Murcia JP, Rethore P-E, Dimitrov N, Natarajan A, Sorensen JD, Graf P, Al. E. Uncertainty propagation through an aeroelastic wind turbine model using polynomial surrogates. *Renewable Energy*. 2017.

18. Werbos PJ. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*: Harvard University; 1975.

19. Jarrett K, Kavukcuoglu K, Ranzato MA, LeCun Y. What is the best multi-stage architecture for object recognition? *Proceedings of the IEEE International Conference on Computer Vision*. 2009:2146-2153.

20. Krizhevsky A, Sutskever I, Hinton GeoffreyE. ImageNet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems 25*; 2012:1097-1105.

21. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. MIT Press; 2016. https://www.deeplearningbook.org

22. Cybenko G. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*. 1989;2(4):303-314.

23. Kingma DP, Ba J. Adam: a method for stochastic optimization. *arXiv e-prints*. December 2014:arXiv:1412.6980.

24. Sudret B, Mai CV. Computing derivative-based global sensitivity measures using polynomial chaos expansions. *Reliability Engineering and System Safety*. 2015;134:241-250.

25. Dimitrov N, Natarajan A, Mann J. Effect of normal and extreme turbulence spectral parameters on wind turbine loads. *Renewable Energy*. 2017;101:1180-1193.

26. IEC. *International Standard IEC61400-1: wind turbines—part 1: design guidelines*, Third; 2005.

27. Zwick D, Muskulus M. The simulation error caused by input loading variability in offshore wind turbine structural analysis. *Wind Energy*. 2015;18(8):1421-1432.

28. Rosenblatt M. Remarks on a multivariate transformation. *Annals of Mathematical Statistics*. 1952;23:470-472.

29. Johannessen K, Meling TS, Haver S. Joint distribution for wind and waves in the northern North Sea. *Proceedings of the Eleventh (2001) International Offshore and Polar Engineering Conference*. 2001;I(1):19-28.

30. Caflisch RE. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*. 1998:1-49.

31. Larsen TJ, Hansen AM. How to HAWC2, the user's manual. R-1597, Technical University of Denmark, Department of Wind Energy; 2012.

32. Hansen MH, Henriksen LC. Basic DTU wind energy controller. E-0028, Technical University of Denmark, Department of Wind Energy; 2013.

33. Kelly M, Larsen G, Natarajan A, Dimitrov N. Probabilistic meteorological characterization for turbine loads. *Journal of Physics: Conference Series (Online)*. 2014;524(1):012076.

34. Keck R-E, de Mare MT, Churchfield MJ, Lee S, Larsen G, Aagaard Madsen H. On atmospheric stability in the dynamic wake meandering model. *Wind Energy*. 2014;17(11):1689-1710.

35. Larsen G, Verelst DR, Bertagnolio F, Ott S, Chougule A. In search for a canonical design ABL stability class for wind farm turbines. *Journal of Physics: Conference Series*. 2016;753(3):032015.

36. Murcia Leon JP, Réthoré P-E, Natarajan A, Sørensen JD. Uncertainty quantification in wind farm flow models. *Ph.D. Thesis*. Denmark; 2017.