https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Multi-task Learning Long Short-term Memory Model to Emulate Wind Turbine Blade Dynamics

Shubham Baisthakur [1] and Breiffni Fitzgerald [1]

[1]School of Engineering, Trinity College Dublin, Ireland

Correspondence: Breiffni Fitzgerald (Breiffni.Fitzgerald@tcd.ie)

Abstract. The high computational costs in the dynamic analysis of wind turbines prohibit efficient design assessments and site-specific performance estimations. This study investigates the suitability of various dimensionality reduction techniques combined with a Long Short-term Memory (LSTM) algorithm to predict turbine responses, addressing computational challenges posed by high-dimensional inflow wind fields and complex time-stepping integration schemes. Feature selection criteria and a multi-stage modelling approach are implemented to arrive at a robust model configuration. Additionally, multi-task learning strategy is implemented which enables the LSTM model to predict multiple target variables simultaneously, eliminating the need for separate models for each target variable. Results demonstrate that this combined approach significantly reduces computational costs while maintaining consistent accuracy across all the target variables, thereby facilitating design feasibility studies and site-specific analyses of wind turbines.

1 Introduction

The wind turbines, in an attempt to maximise energy captures, have grown significantly over the last few decades with their scale seeing unprecedented growth (Roga et al., 2022). The increased scale of wind turbines translates to higher loads, deformations and more accumulated damage. Achieving an efficient design in the presence of these challenges is not a trivial task. Studies focused on efficient controls, advanced technologies, and an improved understanding of wind turbine operations have led to a more efficient operation of wind turbines (Sarkar et al., 2020; Sarkar and Fitzgerald, 2020, 2022; Abbas et al., 2022; Njiri and Söffker, 2016; Sun et al., 2012; Tan et al., 2022; Fitzgerald et al., 2023). In addition, use of reliability and optimisation principles in the design of wind turbines can further improve the design ensuring consistent reliability in the wake of these challenges. However, evaluating multiple possible design combinations to arrive at an optimal solution satisfying multiple constraints requires high computational resources. Optimisation studies can quickly become computationally unfeasible to perform with an increasing number of parameters, hindering the use of optimisation principles in the design of wind turbines. For a wind turbine, the computational demand for optimisation is further compounded due to a large number of uncertain parameters involved, site-specific loading envelopes and the high computational cost of running numerical models. To address these issues, the authors present a methodology to develop a machine-learning-based model to predict the dynamic response of a wind turbine at a fraction of the computational cost of a numerical model.

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Dynamic analysis of wind turbines refers to analysing the structural response subjected to stochastic wind inflow during operation. Wind speeds can fluctuate significantly across the rotor plane within the wind turbine rotors, resulting in varying wind conditions experienced at various points within the rotor area. This effect is more pronounced in turbines with higher rotor diameters. Furthermore, since wind speeds change over time, these spatial points within the rotor area are subjected to temporal variations in wind speed. Therefore, accurate prediction of a wind turbine's response to dynamic forces necessitates a realistic simulation of wind speed variations across its large rotor area and their evolution over time. The complexity of the spatio-temporal wind field introduces a fundamental challenge in wind turbine modelling in terms of a high dimensionality of the input space. TurbSim (Jonkman, 2009), a widely used turbulent wind field simulator, is a key tool for this purpose but is impacted by this dimensionality issue. The challenges arising from the high dimensionality of TurbSim data have been highlighted in many studies (Pereira et al., 2019; Haghi and Crawford, 2021; Bashirzadeh Tabrizi et al., 2019). Some studies have explored the use of surrogate modelling approach (Haghi and Crawford, 2023) and dimensionality reduction techniques (Lataniotis, 2019; Garcke et al., 2017) to alleviate this issue. In this study, the dimensionality reduction approach has been used to extract critical information from a high-dimensional representation of the wind field. Dimensionality reduction has been an active area of research in the domain of surrogate modelling (Hou and Behdinan, 2022), processing speech signals (Markaki and Stylianou, 2008), digital photographs (Van Der Maaten et al., 2009), or medical imagery(Hamarneh et al., 2011). Specialised literature on various dimensionality reduction techniques and their comparative performance on standard datasets is presented by Van Der Maaten et al. (2009). By identifying and retaining the most informative features, dimensionality reduction techniques facilitate efficient analysis of a dataset while reducing computational complexity and improving interpretability. To this end, Principal Component Analysis (PCA) and Discrete Cosine Transform (DCT) are used in this study to arrive at a low-dimensional representation of the inflow wind. Further, building on these extracted features, an LSTM model is developed to capture the temporal dependence between the features of inflow wind and structural response.

Long Short-term Memory (LSTM) Models are a type of recurrent neural network (RNN) with an internal memory state that captures the long-term dependence of the input features on the target variable. LSTM models have been implemented successfully in wind turbines for power forecasting (Banik et al., 2020; Yu et al., 2019; Woo et al., 2018) and damage detection (Choe et al., 2021; Xiang et al., 2021; Chen et al., 2021). Further, Dimitrov and Göçmen (2022) have demonstrated the use of LSTM as a virtual sensor, which can be used to predict the wind turbine parameters which are difficult to measure accurately on-site using SCADA and operational load measurements. Such models are applicable during the operational phase of wind turbines where on-site measurements are available. However, very limited literature exists for the use of LSTMs in structural response prediction of wind turbines during the design and analysis stage where extensive load measurement and SCADA data are not available (Woo et al., 2018; Shi et al., 2023; Zhu et al., 2024; Baisthakur and Fitzgerald, 2024a). In this context, the current manuscript aims to develop an LSTM model for application in the analysis and design stage, focusing on predicting the dynamic response of a wind turbine using the model-generated datasets.

Wind turbines represent a special class of structures whose response is impacted by multiple disciplines including atmospheric modelling, principles of machines, structural dynamics, control engineering and electronics. An efficient surrogate model should be able to integrate the various principles impacting the wind turbine while computing its response. To address

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

this, a multi-stage modelling approach has been used where incremental information about the target is gained in multiple stages. Due to the increasing scale and flexibility of the wind turbines, multiple degrees of freedom are required to model a wind turbine structure and capture its intricate deformation patterns. Therefore, in order to get complete information about the system, the response at each DOF needs to be evaluated. However, creating a surrogate model for each DOF would necessitate developing multiple surrogate models. As the number of DOFs increases, the computational burden associated with training and implementing individual models can increase significantly. To address this, multi-output learning, also known as multi-task learning has been used in this study to model multiple target variables using a single LSTM model (Thrun and Mitchell, 1995; Caruana, 1997). Multi-task learning leverages the inherent relationships between different target variables, to develop a single, unified model capable of simultaneously predicting multiple target variables for a given set of input parameters. This approach is particularly useful in modelling structural response where multiple response variables are closely related to each other and are driven by a common external force. The use of dimensionality reduction techniques with a feature selection algorithm and multi-task learning approach leads to an efficient LSTM model capable of emulating the dynamics of wind turbines. The manuscript is organized as follows: section 2 presents the approach used for generating stochastic wind fields, followed by section 3 which describes the numerical model of the wind turbine used in this study. Section 4 provides a brief description of the dimensionality reduction algorithms used in this study, and section 5 discusses the formulation of the LSTM algorithm and multi-task learning approach. Further, section 6 and section 7 describe the rationale behind selecting the input and output parameters for the surrogate model, respectively. Finally, section 8 presents the numerical results and analyses the model performance in terms of achieved accuracy and computational efficiency. The manuscript concludes with section 9 which summarises the key findings and contributions of this study.

2 Generation of stochastic wind field

Wind turbines are subjected to stochastic turbulent wind inflow during their operation. The power generation and load acting on the components are mainly governed by the properties of the inflow wind. As such, generating a realistic wind inflow pattern is crucial to ensure the validity of simulation results. In this study TurbSim (Jonkman, 2009) is used to generate the wind field acting on the wind turbine. TurbSim serves as a stochastic simulator for turbulent winds, generating full-field representations of wind dynamics. TurbSim operates on a statistical model to produce time series data comprising three-dimensional wind speed vectors across a fixed two-dimensional vertical grid. This grid remains stationary in space throughout the simulation process. To generate wind fields specific to a wind turbine configurations or site conditions, TurbSim provides customizable parameters encompassing turbine geometry and meteorological factors that define the wind speed field. For comprehensive insights into these parameters, readers are referred to TurbSim's user guide (Jonkman, 2009). Here, a summary of key variables used in this study, with their definitions and significance, is presented. These parameters control the size and complexity of the wind datasets:

1. Random Seeds: TurbSim uses two random seeds (*RandSeed1* and *RandSeed2*) to create random phases, one per frequency per grid point per wind component, for the velocity time series and for random number generating schemes.

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

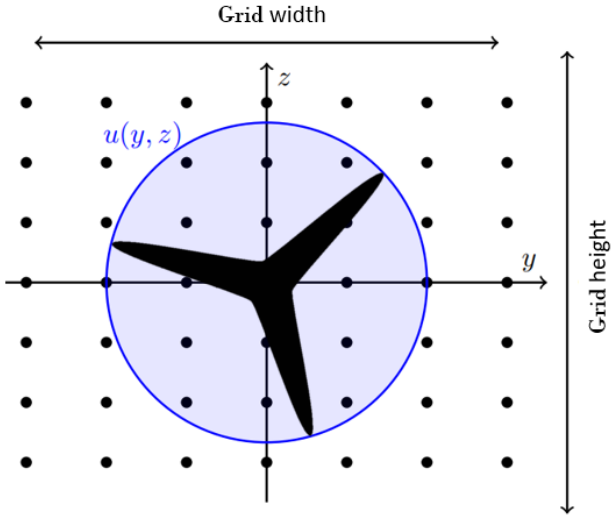


Figure 1. Visual representation of TurbSim Grid

These seeds primarily ensure the reproducibility of the synthetic wind fields and generate unique stochastic realisations for specified environmental conditions. Further, these seeds can be varied at constant meteorological parameters to produce various stochastic patterns with similar statistical properties.

2. Turbine/Model Specification: In configuring TurbSim, the parameters *GridHeight* and *GridWidth* define the total vertical expanse within which the wind turbine rotor operates. This region is then subdivided into segments defined using $NumGrid_Y$ and $NumGrid_Z$, governing the spatial resolution of the computational domain. The variable *HubHt* designates the hub height of the wind turbine and act as a reference point for grid placement. Horizontally, the rotor is centrally aligned within the grid. A representative image of the TurbSim grid encompassing the wind turbine rotor is presented in Fig. 1

3. Meteorological Boundary Conditions: Within TurbSim, the variable *TurbModel* defines the spectral model utilized for generating the wind speed field. The choice of spectral model is guided by site characteristics and the type of analysis. The statistical properties characterizing the inflow, notably mean wind speed, turbulence intensity, surface roughness length, and power law exponents, are defined through variables *Uref*, *IECturbc*, $Z0$, and *PLExp*, respectively. The mean wind speed is specifically defined at a designated reference height, denoted as *RefHt*.

TurbSim employs spatial and temporal discretization techniques to analyse the wind field comprehensively by partitioning the domain into grid cells, enabling the calculation of wind speeds and turbulence properties at each grid point. The temporal dimension of the spatial grid is subdivided into small intervals to accurately capture temporal fluctuations in wind speed and turbulence. Using the specified turbulence model and input parameters, TurbSim generates synthetic turbulent variations in wind speed along temporal and spatial coordinates across the domain. This process involves simulating the spatial and tempo-

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

ral evolution of turbulence based on statistical properties derived from the chosen turbulence model. The IEA-15MW reference wind turbine modelled in this study, characterized by a 240m rotor diameter and a 150m hub height, is modelled using a rectangular domain measuring 285m x 285m, featuring a 25 x 25 spatial grid layout. Considering these turbine specifications, a single 10-minute duration wind field, with a temporal discretization of 0.05 seconds, yields a substantial 625 features with 1201 observations each, leading to 750,625 individual observations for a single wind field. The dimensionality of the input space is further compounded as multiple simulations are required to train a surrogate model. An LSTM model may encounter challenges, such as overfitting and increased computational and memory requirements when dealing with such high-dimensional input data. To mitigate these issues, dimensionality reduction techniques are used to extract essential information from the input space while minimizing the number of variables. Dimensionality reduction facilitates streamlined and efficient model training and implementation. The dimensionality reduction algorithms explored in this study are presented in the section 4.

The wind field generated using TurbSim acts on the wind turbine and govern the load and deformation response. The next section describes the theoretical formulation of the numerical model of wind turbine. This model is further used to generate the data required for training the surrogate model.

3 Numerical Model of the IEA-15MW Wind Turbine

The numerical model of wind turbine used in this study is developed using a multi-body dynamics methodology, based on Kane's dynamics principles (Kane and Levinson, 1985). The Knae's dynamics approach is very effective in managing the intricate interactions among various components of the turbine, facilitating a precise representation of the overall system dynamics. By employing Kane's method, the complexity of deriving equations of motion is significantly reduced, and it allows for a simplified computer implementation compared to traditional methods such as Euler-Lagrange and D'Alembert's principle.

In total, 22 degrees of freedom are included to accurately represent the dynamics of the wind turbine components. The foundation is modelled with six degrees of freedom, incorporating three translational and three rotational motions. The tower is represented using the modal summation technique, which involves four principal mode shapes that characterize the tower's movements in both the fore-aft and side-to-side directions. However, the axial shortening and twisting of the tower due to external loads are not taken into account. To ensure an accurate depiction of rotor speed, the azimuth of the generator and the twisting of the low-speed shaft are included in the model. The blades are treated as flexible elements, employing the modal summation method with three mode shapes for each blade—two modes for flapwise deformations and one mode for edgewise deformations.

Multiple reference frames are established to articulate the motion of different system components and to define their orientations relative to one another. The equilibrium equations for a simple holonomic multi-body system, derived using Kane's approach, are expressed as follows:

$$F_r + F_r^* = 0 \tag{1}$$

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Here, F_r represents the generalized active forces, while F_r^* denotes the inertia force. These forces can be expressed in terms of kinematic variables as follows:

$$F_r = \sum_{i=1}^{n} {}^E v_r^{X_i} \cdot F^{X_i} + {}^E \omega_r^{N_i} \cdot M^{N_i} \tag{2}$$

$$F_r^* = -\sum_{i=1}^{n} {}^E v_r^{X_i} (m^{N_i} {}^E a^{X_i}) - {}^E \omega_r^{N_i} \cdot {}^E \dot{H}^{N_i} \tag{3}$$

In these equations, F^{X_i} is the force vector acting on the center of mass of point X_i, and M^{N_i} is the moment vector acting on the rigid body N_i. The terms ${}^E v_r^{X_i}$ and ${}^E \omega_r^{N_i}$ indicate the partial linear and angular velocities of point X_i and rigid body N_i, respectively. Additionally, ${}^E \dot{H}^{N_i}$ represents the time derivative of the angular momentum of rigid body N_i about its center of mass X_i in the inertial frame, expressed by the equation:

$${}^E \dot{H}^{N_i} = \bar{\bar{I}}^{N_i} \cdot {}^E \alpha^{N_i} + {}^E \omega^{N_i} \times \bar{\bar{I}}^{N_i} \cdot {}^E \omega^{N_i} \tag{4}$$

The final governing equation of the system is structured as follows:

$$M(q,t)\ddot{q} + f(q,\dot{q},t) = 0 \tag{5}$$

In this expression, $M(q,t)$ denotes the inertia matrix, while $\ddot{q}$ is the acceleration vector. The function $f(q,\dot{q},t)$ represents the force vector, which comprises both external and restoring forces acting on the structure. Numerical methods are employed to solve this system of equations, specifically the fourth-order Runge-Kutta method in this study. A comprehensive derivation of these motion equations is outside the scope of this work, but interested readers may refer to Sarkar and Fitzgerald (2021) for further details.

The research presented here utilizes the specifications of the International Energy Agency (IEA) 15MW wind turbine, classified as an International Electrotechnical Commission (IEC) Class 1B direct-drive machine, featuring a rotor diameter of 240 meters (m) and a hub height of 150 m. A representative model of the IEA 15MW wind turbine is illustrated in Figure 2, sourced from the technical report detailing the IEA 15-megawatt reference wind turbine Gaertner et al. (2020), with the operational parameters outlined in Table 1.

This wind turbine is modelled following the Kane's dynamics principles described earlier. The multi-body model established in this research has been validated against the National Renewable Energy Laboratory's (NREL) OpenFast model, a widely recognized open-source framework for simulating wind turbine dynamics. OpenFast has been rigorously benchmarked against experimental data, ensuring it closely reflects the actual performance of wind turbines. This validation process serves to confirm the accuracy of the numerical model developed in this study. The results from this validation are illustrated in Figure 3.

The comparison highlights that the numerical model developed in this study can accurately match the response prediction of OpenFast model establishing the numerical model's accuracy. This model is utilized to generate the data required for training the machine learning model.

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

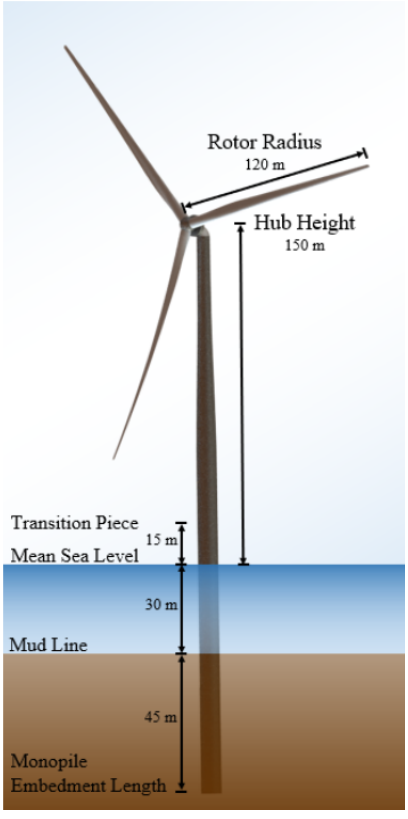


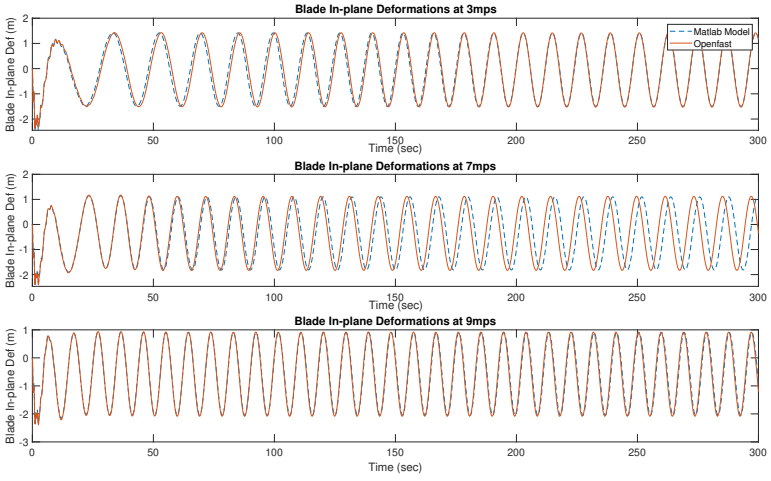
Figure 2. The IEA 15MW reference wind turbine

Parameter	Value
Hub Height	150 m
Rotor Diameter	240 m
Cut-in wind speed	3.00 m/s
Rated wind speed	10.59 m/s
Cut-out wind speed	25.00 m/s
Minimum rotor speed	5.00 rpm
Maximum rotor speed	7.56 rpm

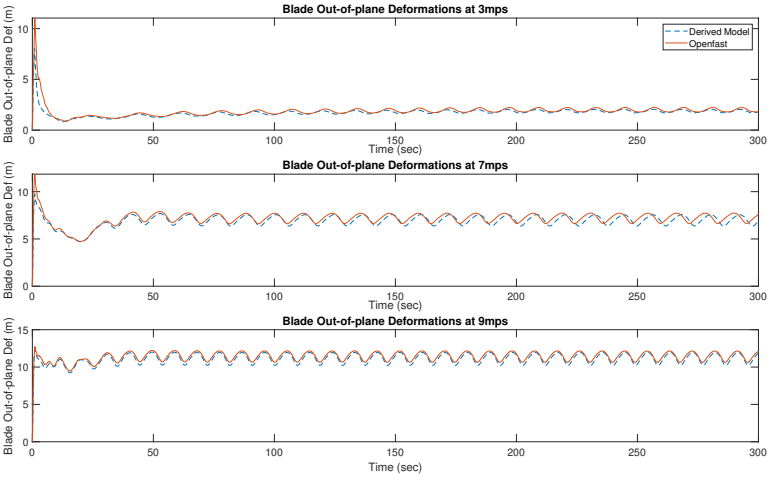
Table 1. Key parameters of the IEA-15MW wind turbine

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS



(a) Blade in-plane deformations



(b) Blade out of plane deformations

Figure 3. Model Verification: Comparison of blade response

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

As highlighted in previous section, the turbulent wind inflow produced using TurbSim, which drives the system response, has a very high-dimensional representation. The next section focuses on reducing the dimensionality of the wind field data using standard dimensionality reduction techniques.

4 Dimensionality Reduction Techniques for Wind Field Data

Dimensionality reduction refers to the process of transforming high-dimensional data into a meaningful representation with reduced dimensionality. Dimensionality reduction is often used as a pre-processing step before building surrogate models for high-dimensional input spaces. The reduced-dimensional representation facilitates faster model training and can improve the robustness of surrogate models.

Mathematically, let $\mathbf{X}$ denote a high-dimensional dataset, where each row corresponds to an observation and each column represents a feature. If p is the number of observations and N is the number of features, then $\mathbf{X}$ is an $p \times N$ matrix. If p is a large number, handling such high-dimensional data poses computational challenges and can lead to inefficiencies in analysis and modelling tasks. Let $\mathbf{Z}$ represent the reduced-dimensional representation of the original dataset $\mathbf{X}$, where $\mathbf{Z}$ is an $p \times n$ matrix with, where $n << N$ being the reduced number of dimensions. The goal of dimensionality reduction is to find a mapping function $f : \mathbf{X} \rightarrow \mathbf{Z}$ that captures the important information in the original data while reducing its dimensionality.

In the context of wind turbines, TurbSim simulations can generate wind fields consisting of thousands of data points across space and time. Dimensionality reduction is crucial to transform these complex wind fields into more manageable representations for efficient response prediction. Various dimensionality reduction techniques are explored in the literature for application to high dimensional problems, of all these techniques, Principal Component Analysis (PCA) is one of the most widely used approach. PCA is a statistical technique which focuses on capturing spatial correlations within the data by identifying principal components that capture the maximum variance in the dataset. PCA offers computational efficiency and linear mapping, making it suitable for handling large-scale wind datasets encountered in wind turbine modelling. Lataniotis (2019) has shown that PCA consistently outperformed the other dimensionality reduction techniques implemented in their research in terms of the reconstruction error achieved, and the robustness of the results over different repetitions on the standard datasets used in dimensionality reduction problems. PCA implemented for wind speed forecasting (Skittides and Früh, 2014; Geng et al., 2020), wind turbine fault detection (Zhang et al., 2021), and monitoring (Wang et al., 2016) has also delivered good results. The principal components can capture recurring spatial and temporal trends in wind speed variations. The reduced-dimensional representation using PCA further helps in data handling and analysis. While PCA excels at capturing the variance within the data through linear relationships, it might overlook the presence of underlying non-linear patterns in complex wind fields. Also, while principal components represent significant variance, interpreting their physical meaning can be challenging for wind fields. To address these potential limitations, the Discrete Cosine Transform (DCT) is used as a complementary technique in this research.

The Discrete Cosine Transform (DCT) offers a powerful tool to analyse wind field dynamics by decomposing the data into its underlying frequency components, revealing patterns of spatial and temporal variation. Unlike PCA, DCT emphasizes

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

the frequency-domain representation of signals or images, leveraging its energy compaction property to highlight dominant frequency components. This data transformation allows DCT to capture the energy content of a signal or image in fewer coefficients by exploiting the frequency-domain characteristics of the data. This property enables effective compression and representation of wind data, which can be exploited for dimensionality reduction by capturing dominant spatial and temporal correlations inherent in wind speed fields. The components obtained through DCT (cosine functions with specific frequencies) directly correspond to spatial variations of different wavelengths or scales within the wind field data. This representation makes the DCT components physically interpretable. Recently, DCT has been applied by Schär et al. (2024) for dimensionality reduction of the stochastic wind field and has been shown to deliver good results.

Given the distinct strengths and limitations of both PCA and DCT for wind field representation, their suitability for reducing the dimensionality of turbulent wind fields is investigated in this study. The mathematical formulation of these methods is presented in the next section.

4.1 Principal Component Analysis

PCA is a mathematical tool that transforms potentially correlated features into a smaller set of uncorrelated variables called principal components. This transformation maximizes the variance explained by each component, thereby emphasizing the most prominent patterns within the data. In PCA the data is normalized to have zero mean to ensures that principal components capture variations from the average behaviour and not the absolute magnitude of the wind speeds. Following the original notations, assume ($\mathbf{X} \in \mathbb{R}^{p \times N}$) represent the original wind field dataset matrix, with p observations (time steps) and N features (spatial locations). The mean-centred dataset ($\mathbf{X'} \in \mathbb{R}^{p \times N}$) is given by:

$$\mathbf{X'} = \mathbf{X} - \overline{\mathbf{x}} \tag{6}$$

where $\overline{\mathbf{x}}$ is the mean vector of the dataset. PCA computes the covariance matrix to quantify the pairwise linear relationships between different features within the wind field data. The covariance matrix ($\mathbf{\Sigma}$) of the mean centred data is computed as:

$$\mathbf{\Sigma} = \frac{\mathbf{1}}{\mathbf{p-1}}(\mathbf{X'^T X'}) \tag{7}$$

Further, eigen-decomposition of the covariance matrix ($\mathbf{\Sigma}$) is performed to identify the directions of maximum variance within the wind field data and quantify the amount of variation captured along each direction. The eigenvectors ($\mathbf{V} \in \mathbb{R}^{N \times N}$ and their corresponding eigenvalues ($\mathbf{\Lambda} \in \mathbb{R}^{1 \times N}$ are represented using the eigenvector matrix $\mathbf{V}$ with N columns and eigenvalue vector $\mathbf{\Lambda}$ with N entries:

$$\mathbf{V} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_N \end{pmatrix}$$

$$\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$$

In PCA the eigenvectors are sorted in decreasing order of their corresponding eigenvalues such that the eigenvector with the largest eigenvalue represents the principal component explaining the most variance in the wind field data. The number of

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

principal components required to capture the maximum information about the wind field data is examined using the cumulative explained variance calculated as:

$$e_k = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{N} \lambda_i} \tag{8}$$

A common approach is to select the smallest number of principal components (k) that achieve a desired percentage of explained variance. Finally, the original high-dimensional dataset $\mathbf{X}$ is projected onto the new basis spanned by the selected principal components to obtain the reduced-dimensional representation $\mathbf{Z} \in \mathbb{R}^{p \times n}$, where n is the desired number of principal components to retain. The projection can be expressed as:

$$\mathbf{Z} = \mathbf{X}'\mathbf{V}_n$$

where $\mathbf{V}_n$ is the matrix containing the first n eigenvectors corresponding to the largest n eigenvalues. This projection transforms the original features representing wind speeds at different grid points into a reduced set of features. By selecting the n principal components with the largest eigenvalues, PCA ensures that these new features capture most of the essential variation within the wind field data.

4.2 Discrete Cosine Transform

The Discrete Cosine Transform (DCT) is a frequency-domain transformation technique which uses real-valued cosine functions as its basis. For a discrete signal $x(n)$ of length N, the 1D DCT is calculated using the Eq. 9

$$z(k) = \sum_{n=0}^{N-1} x(n) \cdot \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right) \tag{9}$$

where $z(k)$ is the k-th frequency component of the signal transformed using DCT. Following the approach implemented by (Schär et al., 2024), a mathematical correspondence can be established between a 2D spatial wind field grid generated using TurbSim and an image. Assuming a wind field defined over a discrete grid of size $(N_x \times N_y)$, each grid point (i, j), representing a spatial location in the wind field, can be mapped to a pixel in the image. The wind speed magnitude at a grid point, denoted by $(X(i,j))$, determines the intensity of its corresponding pixel (see Fig. 4). Using this representation of the wind field, 2D DCT is used to analyse wind speed variations across spatial scales and directions. The 2D DCT of the wind field (X), denoted by $Z(u,v)$, is calculated using the Eq. 10:

$$Z(u,v) = \alpha(u)\alpha(v) \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} X(i,j) \cos\left(\frac{\pi(2i+1)u}{2N_x}\right) \cos\left(\frac{\pi(2j+1)v}{2N_y}\right) \tag{10}$$

such that,

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{N_x}}, & i = 0 \\ \sqrt{\frac{2}{N_x}}, & 1 \le i \le N_x - 1 \end{cases} \tag{11}$$

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

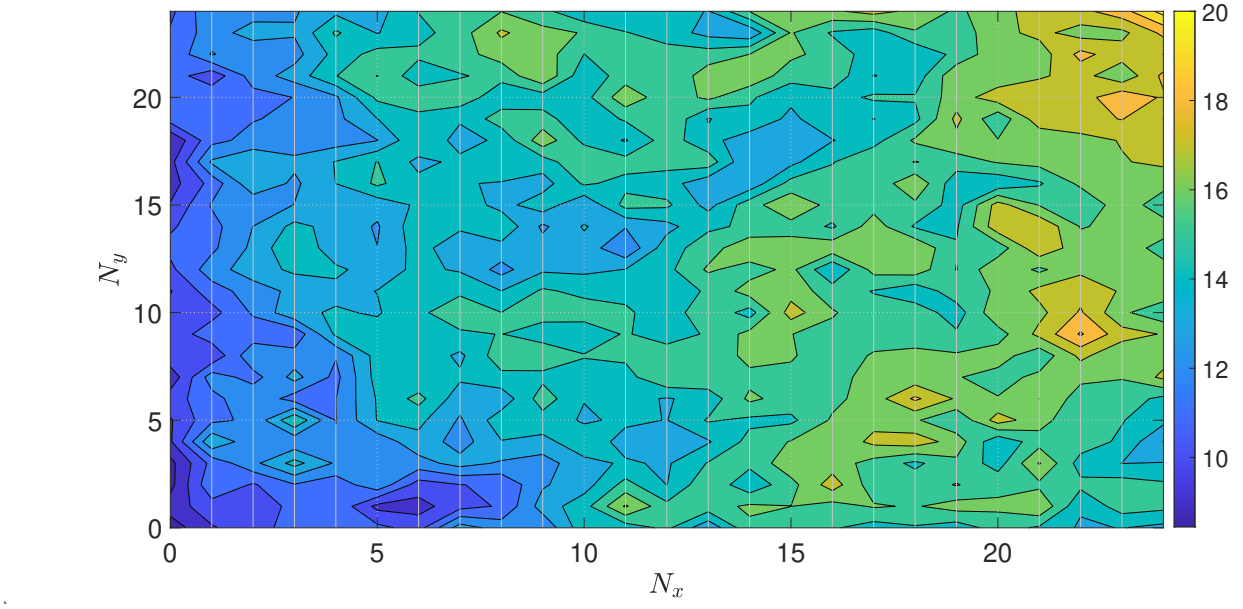


Figure 4. Modelling wind speed data as image

and

$$\alpha(v) = \begin{cases} \frac{1}{\sqrt{N_y}}, & j = 0 \\ \sqrt{\frac{2}{N_y}}, & 1 \leq j \leq N_y - 1 \end{cases} \tag{12}$$

where u and v are frequency indices representing spatial frequencies in the transformed domain and $\alpha(u)$ and $\alpha(v)$ are normalization factors ensuring orthogonality of the basis functions. The coefficients $(Z(u,v))$ obtained from the 2D DCT transformation represent the strength of different frequency components present within the wind field. These frequency components correspond to variations in wind speed at different spatial scales. Low-frequency coefficients, associated with low values of u and v, represent large-scale spatial patterns or trends across the wind field, which captures smooth variations over spatial domains in the wind field data. Conversely, high-frequency coefficients represent finer-scale, localized wind speed fluctuations and turbulent structures. These coefficients capture rapid, often less spatially organized changes in wind speed over short distances within the wind field. By analyzing the energy distribution across different DCT coefficients, the dominant spatial scales of variation present in the wind field can be identified. Using the 2D DCT of a wind field and ranking the coefficients by their magnitudes, a subset of coefficients that captures the majority of the wind field's essential variations can be identified. Retaining the significant coefficients and discarding the high-frequency components results in a compressed representation of the wind field, preserving the dominant spatial patterns in the wind field. Further, an LSTM model is developed using these

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

features to predict the deformation response of wind turbine blades. The details of LSTM model architecture are presented in the next section.

5 Long Short-term Memory Model and Multi-task learning

An LSTM model is a type of Recurrent Neural Networks (RNN) designed to work with sequential data and capture the long-term and short-term dependencies between the input and output variables. An RNN processes sequential data by maintaining the memory of previous inputs through an internal state. RNNs employ feedback mechanisms, where the network's output at each step is fed back as input. This recursive nature enables RNNs to capture temporal dependencies inherent in sequential data. Traditionally, the RNNs suffer from vanishing gradients where the algorithm struggles to preserve the internal state for long-range sequences (Hu et al., 2018). This phenomenon, known as a vanishing gradient, where the gradient of the loss function with respect to the learnable parameters becomes so small that it results in slow or ineffective learning. The LSTM architecture alleviates this issue by using various gates to control the flow of information (Hochreiter and Schmidhuber, 1997). These gates control the flow of gradients during backpropagation, allowing LSTM networks to propagate gradients over long sequences and through multiple layers effectively. An LSTM unit consisting of all these features is known as a memory cell. Selective data retention and elimination are performed as the data passes through these gates. Based on the task of each gate, they are commonly known as:

- Input gate: This gate receives the model features at the current time step and hidden state predicted at the previous time step as the input. This gate combines these two inputs to create a candidate for storing new information. Mathematically, this operation is represented in Eq. 13

$$i_t = \sigma_g(W_i \times x_t + R_i \times h_{t-1} + b_i) \tag{13}$$

- Forget gate: The input to the forget gate is the same as the input gate. This gate combines the information from input features at the current time step with the hidden state from the previous time step and estimates what past information is no longer relevant. This operation is presented in Eq.14

$$f_t = \sigma_g(W_f \times x_t + R_f \times h_{t-1} + b_f) \tag{14}$$

- Cell state: The cell state in an LSTM unit undergoes updates influenced by the outputs of the forget gate and the input gate, which determine the significance of new and existing memory. A cell state is the actual memory of the LSTM, which holds the information across many time steps. The cell state acts as the long-term memory of the unit. This operation is represented through Eq. 15 and Eq. 16

$$c'_t = \sigma_g(W_c \times x_t + R_c \times h_{t-1} + b_c) \tag{15}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot c'_t \tag{16}$$

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

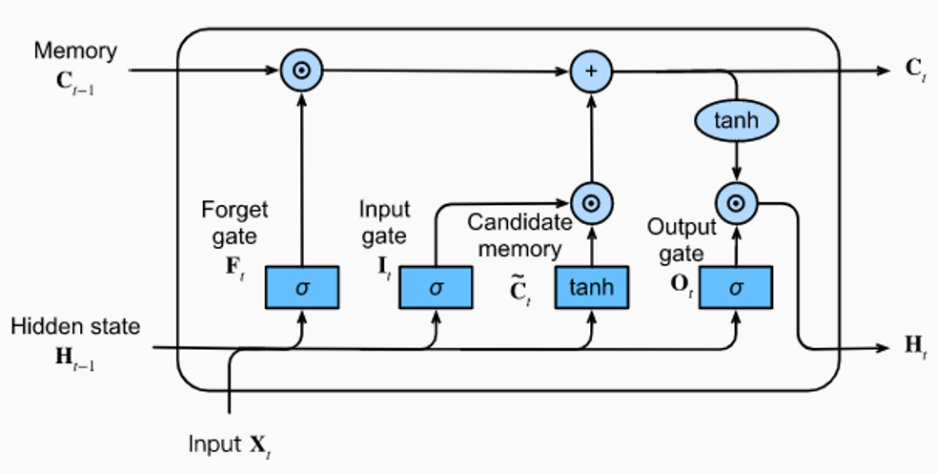


Figure 5. LSTM Architecture

– Output gate: The output gate combines the cell state updated at the current time step with the previous hidden state, creating the current hidden state predicting the network's output (hidden state) at the current time step. This gate acts as the short-term memory of the unit. Eq. 17 and Eq. 18 represents this operation mathematically

$$o_t = \sigma_g(W_o \times x_t + R_o \times h_{t-1} + b_o) \tag{17}$$

$$h_t = o_t \cdot \sigma_c(c_t) \tag{18}$$

In Eqs. 13 through 18, x represents the feature vector of the LSTM model, while h denotes the hidden state, a crucial output of the LSTM unit. The subscript t denotes the temporal position of the vectors, where t refers to the current time step, $t-1$ refers to the previous step, and so forth. The symbols i, f, c, and o correspond to the input gate, forget gate, cell state, and output gate, respectively. In each gate, W signifies the fixed weights, R denotes the recurrent weights, and b represents the bias term. The σ symbol signifies the sigmoid activation function, ensuring non-linearity in the input-output transformation and constraining the gate's output within the interval [0,1]. The sigmoid activation function is given by Eq. 19

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{19}$$

Multiple LSTM units can be stacked together to capture complex relationships in the data. A sketch of typical LSTM memory cell and data flow is presented in the Fig 5, adopted from Ref. (Calazone, 2022).

Multi-task learning is an approach in machine learning where a model leverages shared information between different output variables to learn the dependencies to predict multiple outputs using a set of common input parameters (Caruana, 1997). The multi-task learning model have found to deliver better by jointly learning various dependent parameters than learning them independently (Zhang and Yang, 2018). In multi-task learning method, various deep learning layers are stacked together to

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

efficiently learn patterns from the data. In a standard machine learning approaches, only one output parameter is learned at once; however, multiple parameters equalling the total number of neurons in the final layer can be predicted using a multi-task learning model. An overview of multi-task learning, their application and detailed classification can be found in Ref.(Zhang and Yang, 2018).

6 Data Generation: Selection of Input Variables

A surrogate model learns the relationship between input and output purely based on the data passed during the training process, therefore the generation of a representative dataset covering the various possible scenarios that can be encountered on-site is an important step in the model development. To this end, selection of wind field parameters and their distributions is crucial, as these define the conditions for which the surrogate model is applicable. While the surrogate model can extrapolate results beyond the training dataset bounds, its prediction accuracy in those regions is uncertain. This study focuses on onshore wind turbine performing in normal operating conditions. This load case is chosen as it significantly influence fatigue loading and power production and analyzing these frequent conditions for statistically significant assessment requires running a large number of simulations, which leads to significant computational costs. Developing a surrogate model for these conditions allows for a computationally efficient fatigue and energy yield assessment calculations.

The inflow wind patterns encountered during operation are impacted by a multitude of parameters, however modelling the impact of all these parameters is computationally prohibitive. For the development of surrogate model, it is important to identify a set of parameters which significantly impact the target variables. researchers have established that the turbulence intensity (*TI*), mean wind speed (U), and power law exponent (α) have the highest impact on wind turbine loads under normal operating conditions. Dimitrov et al. (2018) performed a sensitivity analysis of environmental factors with focus on fatigue and extreme loading, and found that mean wind speed and turbulence intensity have the maximum impact while parameters like wind direction, turbulence length scale, and air density exhibit minimal influence within the parameter bounds considered in their analysis. Further, as this study aims to accurately model the blade response, the inclusion of the power law exponent (α) is particularly important as wind shear induces cyclic changes in the wind velocity experienced by the rotating blades, significantly influencing blade root loads. To this end, the *TI*, U and α are selected for generating the wind fields. The parameter bounds used for generating the wind speed data are presented in Table 2. Using these parameter bounds along with the Sobol sampling method, a total of 50 wind speeds ranging from cut-in to cut-out wind speeds are generated. For each wind speed multiple *TI* and α values are generated.

The spatial correlations, temporal variability, and turbulence inherent in wind make it a fundamentally random process. This randomness stems from aleatory uncertainty due to the inherent variability of wind speed and epistemic uncertainty arising from limitations to model or measure all factors influencing wind behaviour. To train a surrogate model capable of accurately predicting wind turbine responses under these uncertainties necessitates the generation of multiple realizations (ensembles) for each set of environmental parameters. Quantifying these uncertainties is essential for making realistic risk assessments. However, identifying the minimum number of seeds required so that the surrogate model can efficiently identify

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Parameter	Lower bound	Upper bound
Mean wind speed	3	25
Turbulence intensity	$0.025 \cdot U$	$0.18\left(6.8 + 0.75 \cdot U + 3\left(\frac{10}{U}\right)^2\right)$
Shear exponent	0	$\alpha_{ref,UB} + 0.4\left(\frac{R}{z}\right)\left(\frac{U_{max}}{U}\right)$

Table 2. Environmental parameters used and their limits

the underlying patterns in these wind fields is a crucial factor governing the amount of data required for training these models. In a different application, the IEC guidelines (IEC, 2019) recommend considering six random realizations for each mean wind speed for computing fatigue damage, implying that six seeds can effectively capture the variability in wind fields. Further, these guidelines recommend keeping environmental conditions constant within a wind speed bin. This approach, combined with the inherent variability in *TI* and α, leads to a large number of required simulations. A study by Hübler et al. (2018) demonstrates that assuming constant environmental conditions within a wind speed bin does not fully capture the uncertainty in fatigue damage and recommend scattering environmental conditions within a bin for a more comprehensive assessment. A combination of these sampling approaches is used in this study to simulate the wind fields required for training the surrogate model. Scattering environmental conditions within a wind speed bin also acts as a crucial data augmentation technique for the surrogate model, which prevents overfitting to a limited set of wind patterns and significantly improves generalization. To this end, a total of 15 realizations are generated for each mean wind speed U in this study, assuming a random combination of *RandSeed1*, *TI*, and α. This ensures a broader sampling of wind patterns while considering computational constraints. As such, wind fields are generated using TurbSim with the following key settings:

1. Spectral Model: Kaimal spectral model (IECKAI)
2. Turbulence Model: Normal Turbulence Model (NTM)
3. Grid Size: 25 x 25 grid with a spatial extent of 285m x 285m
4. Temporal Discretization: 0.05-second

These parameter choices ensure the generated wind fields match the scale and characteristics relevant to the IEA-15MW wind turbine model being studied. The response of wind turbine for these loading conditions is simulated using the numerical model presented in section 3 with a time step of 0.005-second to ensure accuracy and numerical stability. In this study, the ROSCO controller (Abbas et al., 2021, 2022) is used to optimise wind turbine performance by managing rotor speed and blade pitch angle. The structural response is then downsampled to align with the temporal resolution of the wind field data. The wind field is represented by a 25×25 grid over a 600-second duration, resulting in a large input matrix 12000×625. This high dimensionality underscores the importance of dimensionality reduction techniques. The PCA and DCT operations transform a wind field grid into different representations, providing a different basis for capturing the spatio-temporal variations, however, it does not directly reduce the number of features. To this end, the methodology presented in Baisthakur and Fitzgerald (2024b)

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

is adopted for feature selection for multivariate time series problem. Following this approach, Recursive Feature Addition (RFA) method is used to select an optimal subset of features from both the PCA and DCT representations. The RFA iterative techniques operate by systematically adding the features to the input space, to identify an optimal subset of input features that maximize the predictive performance of the machine learning model. RFA starts with an empty feature set and iteratively adds features based on their individual relevance to the target variable. The relevance of a feature is assessed by evaluating the impact of each feature on a chosen performance metric, such as the root mean squared error (RMSE) employed in this study. The feature leading to the lowest RMSE is then added to the input feature set. This process continues by evaluating all remaining features in combination with the chosen ones, exploring the combination that further enhances the model performance. These steps are repeated iteratively until a pre-defined stopping criterion is met, such as reaching a desired number of features or achieving a satisfactory level of model performance. The method terminates when adding extra features does not improve the model performance. Interested readers are referred to Baisthakur and Fitzgerald (2024b) for implementation of RFA method and its analysis of its accuracy and computational cost as compared to other feature selection methods used for multi-variate time series problems. PCA and DCT dimensionality techniques, in combination with the RFA feature selection algorithm, are employed to reduce the feature space and identify the most influential parameters.

7 Selecting the output parameter

A surrogate model's ability to provide comprehensive insights depends on the selected output parameters. In literature various surrogate models are developed to predict quantities which pertain to load components, such as fatigue damage or reactions in the tower and blades (Haghi and Crawford, 2023; Schär et al., 2024; de N Santos et al., 2023; Bai et al., 2023). However, these load components are derived using responses at individual DOFs and the mapping from DOF response to load reactions is straightforward and does not necessitate extensive computational resources. In this context, the authors aims to predict dynamic response at each blade degrees of freedom (DOF). Further, to have physicaly measurable results, another LSTM model is developed to predict the blade in-plane and out-of-plane deformation response. This choice of target variables ensures that the surrogate model is versatile enough to compute multiple derived quantities.

In this approach, multi-task learning streamlines the modelling process and enhances the model's capacity to capture complex interactions and dependencies within the system by exploiting shared information across related outputs. In the context of wind turbine blade analysis, the responses of different DOFs are often governed by common load elements acting on the blade structure. By jointly modelling these responses within a multi-task learning framework, the model can leverage shared features and patterns, leading to improved generalization and predictive performance. The numerical model developed in this study characterize the blades using three DOFs: q_{B1F1}, q_{B1F2} and q_{B1E1}. The multi-output learning paradigm is employed to develop an LSTM model to forecast the time history responses of these three DOFs. Further, the model is extended to predict the in-plane and out-of-plane blade deformations. The results obtained through this approach are presented in the next section.

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

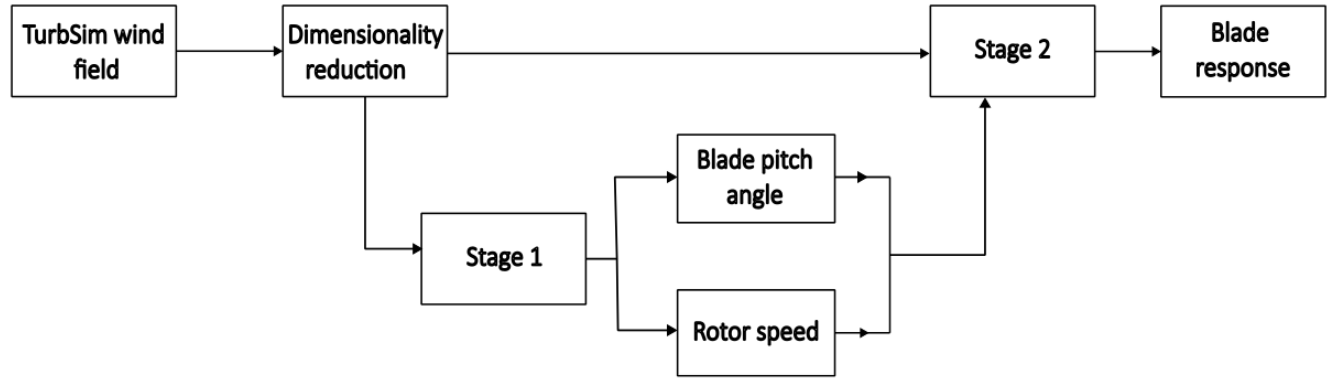


Figure 6. Multi-stage modelling approach for predicting blade response

8 Numerical Results

In this section numerical results demonstrating the performance of LSTM models in predicting the dynamic response of wind turbine blades under turbulent wind conditions are presented. In the initial stage of the analysis, the models were developed to predict blade deflection responses using the reduced order representation of wind field obtained using PCA and DCT techniques. While the models trained using PCA and DCT components as input features could capture the time varying mean of blade response, these models struggled to model fluctuations around the mean responses. This limitation stems from the choice of input features, which lacks direct information about the control parameters (rotor speed and blade pitch angle), which significantly impact the dynamics of blade response.

To improve the prediction accuracy of the LSTM model, a multi-stage modelling approach is implemented. In this approach, the control features i.e., pitch angle and rotor speed, are first modelled as a function of PCA and DCT components. In the second stage, the control parameter predictions are combined with the reduced order representation of wind field to predict the blade response. In essence, the multi-stage modelling approach divides the model development into smaller sub-tasks, which provide incremental information about the final output through a series of sub-tasks. This concept is demonstrated in Fig. 6. A similar multi-stage modelling approach is implemented by (Schär et al., 2024) for predicting wind turbine response using autoregressive models on manifolds. The results for the multi-stage modelling approach are presented in the next section.

8.1 Predicting Control Parameters

This section presents the methodology used to develop the LSTM model for predicting the control parameters. The numerical results are presented and model performance is comprehensively analyzed. The control parameters i.e., rotor speed and blade pitch angle are governed by the properties of the inflow wind. These parameters are tuned using a specially designed controller

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

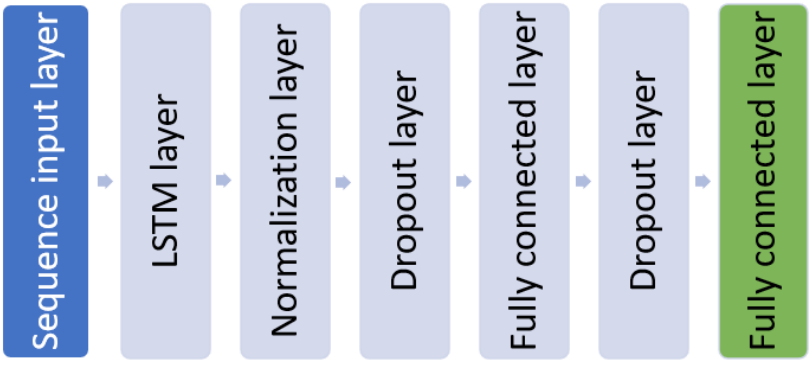


Figure 7. LSTM Model architecture for predicting rotor speed

to efficiently regularize the performance of a wind turbine. The controller used to regularize the rotor speed is aimed at achieving higher efficiency in power production in below-rated wind speeds whereas the blade pitch angle controller aims to maintain constant power production for above-rated wind speeds. The controller algorithm introduces additional uncertainty in the wind turbine response. In this study, the controller algorithm is treated as a black box to develop a surrogate model. Here, an LSTM model is developed to predict the controller response as a function of wind speed data processed through PCA and DCT algorithms. In preliminary analysis, a multi-task learning model was used to simultaneously predict the blade pitch angle and rotor speed as a function of PCA and DCT features. However, this approach failed to simultaneously model both the parameters with the required level of accuracy. Alternatively, individual LTSM models developed to predict one parameter at a time were found to deliver better results. Based on these observations, blade pitch angle and rotor speed are modelled separately using individual LSTM models. Further, it was observed that rotor speed can be better modelled through PCA features and blade pitch angle dynamics are more accurately captured using DCT features. The following sections present the model architecture and its performance in predicting the individual controller response.

8.1.1 Predicting rotor speed response using PCA

In the preliminary analysis, the PCA approach was found to be more effective in modelling the rotor speed response. Supplementing the principal component with rotor-averaged wind speed was found to further improve the model predictions. This section presents the development of the LSTM model trained using principal components of wind speed data and the rotor-averaged wind speed as the input features. The model architecture is shown in Fig. 7. In this deep-learning model, the sequence input layer inputs sequential data to the network, where the size of the sequence layer is equal to the number of input features. The LSTM layer is used to learn long-term dependencies between time steps and extract the temporal patterns from sequential data. This layer performs additive interactions, which can help improve gradient flow over long sequences during training (Hochreiter and Schmidhuber, 1997). Since the model should generalize well over a wide range of wind speeds, a

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Hyperparameter	Parameter range	Optimised values
LSTM layer - 1	[1 - 100]	73
Fully connected layer - 1	[1-50]	50

Table 3. Summary of hyperparameter exploration for predicting controller response

normalization layer was used, which normalizes a mini-batch of data across all channels for each observation independently. The use of the normalization layer after the learning layers speeds up training and improves the model performance. A fully connected layer was then used to combine the temporal patterns and transform these extracted features. While a fully connected layer extracts the learned information from the LSTM layers, it also increases the risk of overfitting. To avoid this, a dropout layer was added to the model. A dropout layer randomly sets input elements to zero with a given probability, which helps avoid overfitting the training data and improves the generalization ability of the network. A constant dropout probability of 10% was assumed for each dropout layer in this architecture. The first fully connected layer was used to aggregate the information from the LSTM layer, while the second fully connected layer was used to represent actual output parameters. The final fully connected layer in the model matches the number of target variables. The final fully connected layer in the LSTM model developed to predict rotor speed has only one neuron. The number of neurons in the final layer of a multi-task learning model is always greater than one. This model was trained using the Adam optimisation algorithm (Kingma and Ba, 2014). Initially, the first five principal components, cumulatively capturing 50% of the variance of wind speed data, were selected to capture the dominant patterns in wind speed data. This feature set was further filtered using RFA by iteratively adding features and evaluating model performance to find the optimum feature set that yields the best results. During the RFA process, rotor averaged wind speed was considered as a fixed element in the input set. Table 3 summarizes the range of hyperparameters used for exploration in this study. The hyperparameters are tuned using Bayesian optimisation technique, the optimised parameter values are also presented in Table 3.

Fig. 8 shows the RMSE obtained from validation data using different principal components as input features. This figure shows that the LSTM model developed using the first principal component combined with rotor average wind speed corresponds to the lowest prediction error on the validation dataset. Adding more input features using the RFA method did not increase the accuracy of model predictions. To this end, the first principal component of the spatio-temporal wind along with rotor averaged wind speed are used as input features to predict the rotor speed. A visual representation of the LSTM model predictions to the actual rotor speed across different wind fields is presented in Fig. 9, demonstrating the accuracy of the model. The overall efficiency of this approach can be contextualized through the fact that the proposed approach reduces 625 input features to only two variables capable of capturing the overall dynamics. This underscores the importance of dimensionality reduction techniques like PCA in simplifying complex datasets without sacrificing crucial information.

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

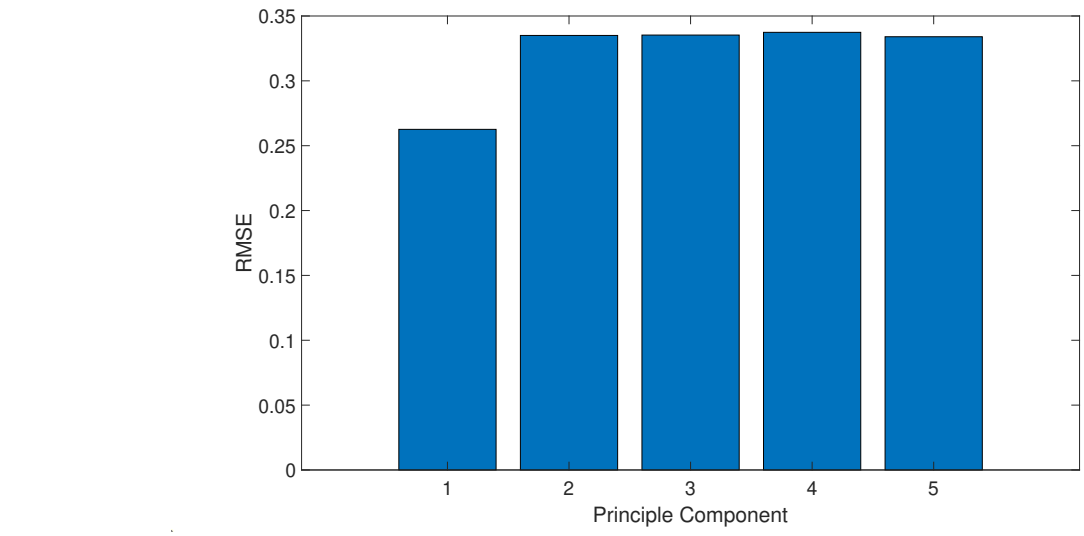


Figure 8. RMSE in rotor speed prediction using different principal components

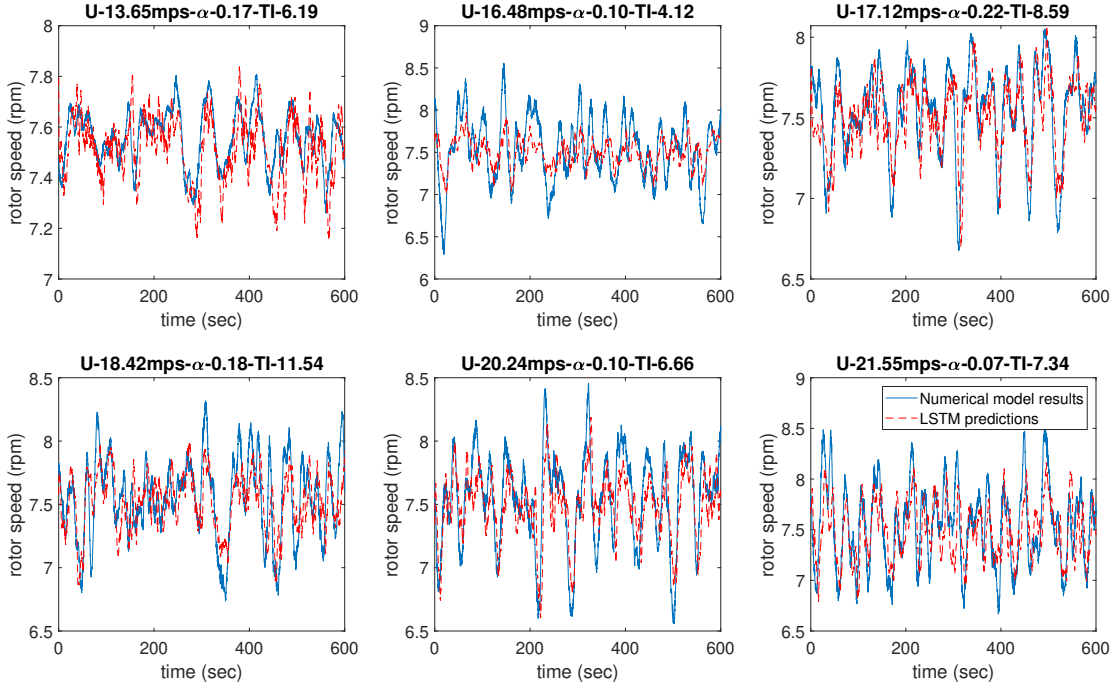


Figure 9. Rotor speed predictions of LSTM model using PCA feature

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

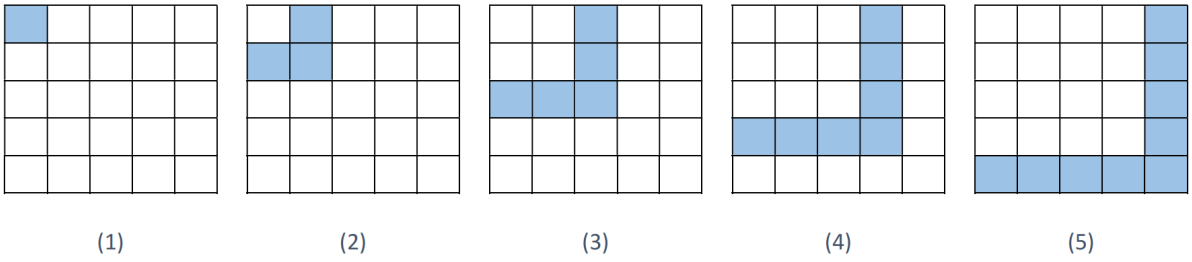


Figure 10. Visual representation of DCT components considered in RFA method

8.1.2 Predicting blade pitch response using DCT

The LSTM model used for predicting blade pitch response shares the same architecture as shown in Fig. 7. Wind speed data, transformed using DCT, served as the input, and RFA was used to determine the most informative subset of DCT features for optimal prediction accuracy. Similar to the rotor speed prediction model, rotor-averaged wind speed acts as a fixed element of input feature set during the RFA process of the blade pitch prediction model. In this model, the wind speed grid at each time step was treated as an image. 2D DCT is applied to this image to decomposes it into different spatial frequency components. The process of modelling wind speed data as an image and processing it through 2D DCT inherently distributes relevant information across the off-diagonal terms of the output matrix. As such, the off-diagonal terms in the DCT matrix share mutual information about the spatial frequencies in the horizontal and vertical directions. To capture this information, the DCT components were grouped to form feature subsets, each representing an incremental higher frequency subset. A mathematical definition of a feature subset z_i is given by Eq. 20

$$z_i = Z_{n\times n}(1:i,1:i) - Z_{n\times n}(1:i-1,1:i-1) \forall 1 \leq i \leq n \tag{20}$$

A visual representation of component groups for a 5×5 DCT matrix is shown in Fig. 10. Using this feature definition along with the RFA scheme, the LSTM model was trained to predict the blade pitch response.

The initial 5×5 sub-matrix of the 25×25 matrix obtained through 2D DCT was selected as the dominant subset of input features as it represents the lower frequency patterns which can more effectively capture the general trends in the wind speed data. Fig. 11 presents the RMSE obtained for the LSTM model trained on each DCT component group (presented in Fig. 10), demonstrating that the first component, representing the lowest frequency patterns extracted by the DCT, achieves the best performance in predicting blade pitch angle. This signifies that low-frequency fluctuations in wind speed play a dominant role in driving the blade pitch response. The findings of the RFA showing the best predictive ability of the LSTM under the low-frequency component can be explained using the control theory. The control system governing blade pitch angle is designed to optimise turbine performance and stability by filtering out high-frequency disturbances and adjusting the blade pitch angle in a more deliberate and controlled manner. This intentional smoothing of blade pitch variations serves to mitigate the impact of sudden changes in wind speed and maintain the overall stability and efficiency of the turbine operation. Fig. 12 further

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

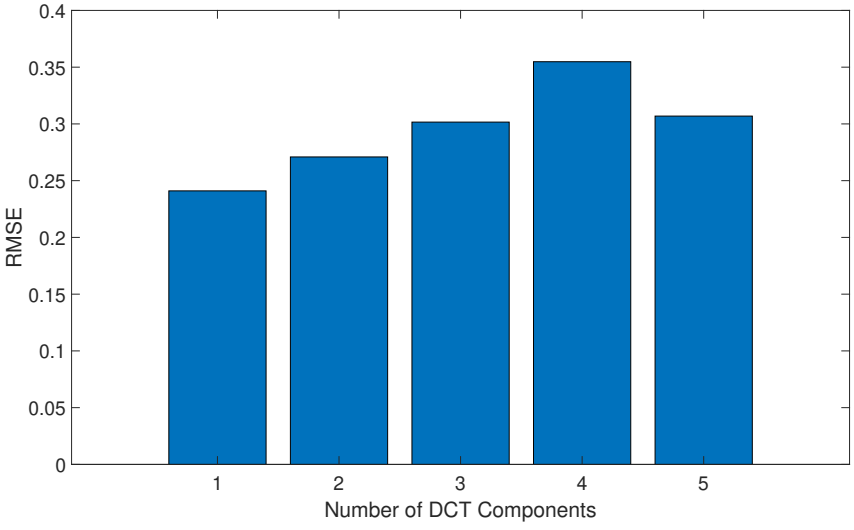


Figure 11. RMSE in blade pitch prediction using different DCT components

reinforces this observation by comparing the actual blade pitch response to the LSTM model's predictions for different wind fields. A good agreement in the blade pitch response across different mean wind speeds showcases the generalization ability of the LSTM model. The rotor speed and blade pitch angle predictions are combined with the PCA and DCT data to predict the dynamic response at the blade tip. In this stage, the rotor speed was transformed into the angular position of the blade so that it could more effectively capture the frequency of the rotation.

8.2 Predicting blade response

In this section, a two-layer LSTM model architecture was implemented to predict wind turbine blade response. To explore the effectiveness of different dimensionality reduction techniques, two sets of input features are considered:

1. Wind turbine data processed through PCA, combined with the angular position of the blade and blade pitch angle predictions from the LSTM models described in Sections 8.1.1 and 8.1.2.

2. Wind speed data processed through DCT, also combined with the angular position of the blade and blade pitch angle predictions from the LSTM model.

The results from these two approaches were compared to assess the capabilities of PCA and DCT in the context of blade response prediction. A multi-task learning approach was employed to simultaneously predict the blade response. Initially, the model was trained to predict the in-plane and out-of-plane deflections, which was subsequently extended to predict the response of individual blade DOFs in flapwise and edgewise directions. Matlab®codes used for model training and validation are available at Baisthakur (2024).

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

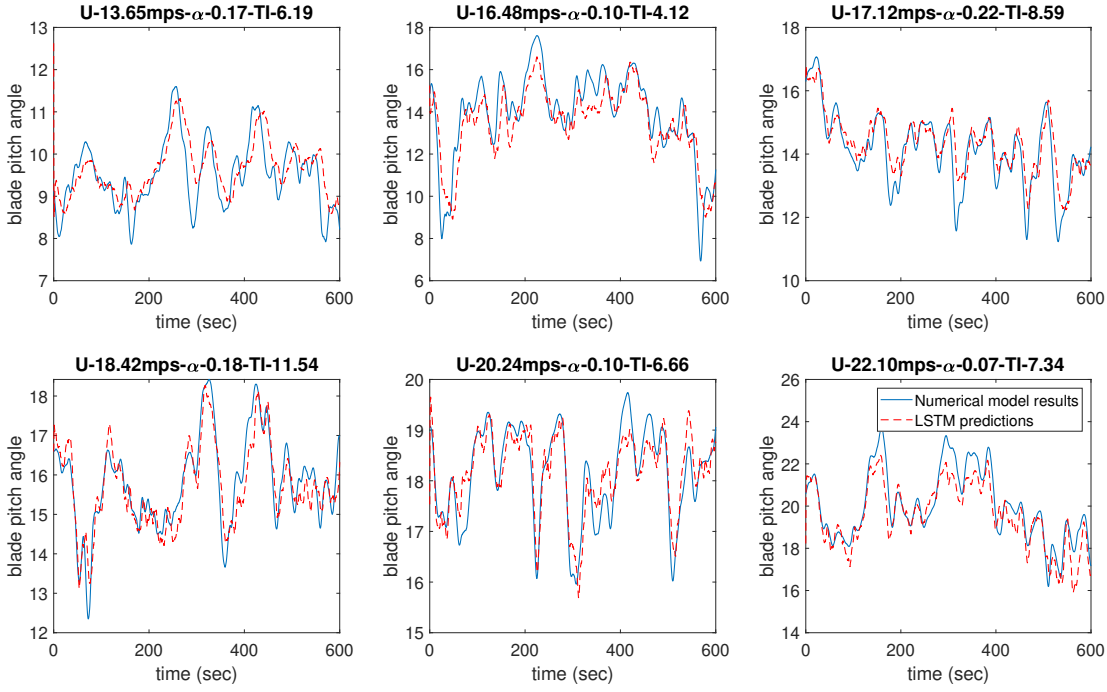


Figure 12. Blade pitch angle predictions of LSTM model using DCT features

Hyperparameter	Parameter range	Optimised values
LSTM layer - 1	[100 - 200]	138
LSTM layer - 2	[50 -100]	63
Fully connected layer - 1	[1-50]	44

Table 4. Summary of hyperparameter exploration to optimise the number of hidden units in learning layers

8.2.1 Blade response prediction with PCA features

Following the approach outlined in the previous sections, this section presents the model developed for predicting blade response and corresponding results. The architecture of the model for blade response prediction is detailed in Fig. 13. The LSTM architecture is designed through an iterative process, and hyperparameter optimisation is used to identify the most optimal parameters for various layers. The number of hidden units in learning layers was used as the hyperparameter in this study. Model performance is evaluated using RMSE between predictions on validation data and the actual results. Table 4 summarizes the hyperparameter exploration. A progressively smaller range is used for the bounds of hyperparameters, as consistently decreasing the number of hidden units was found to deliver better results. Using the angular position of the blade and blade pitch angle as

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

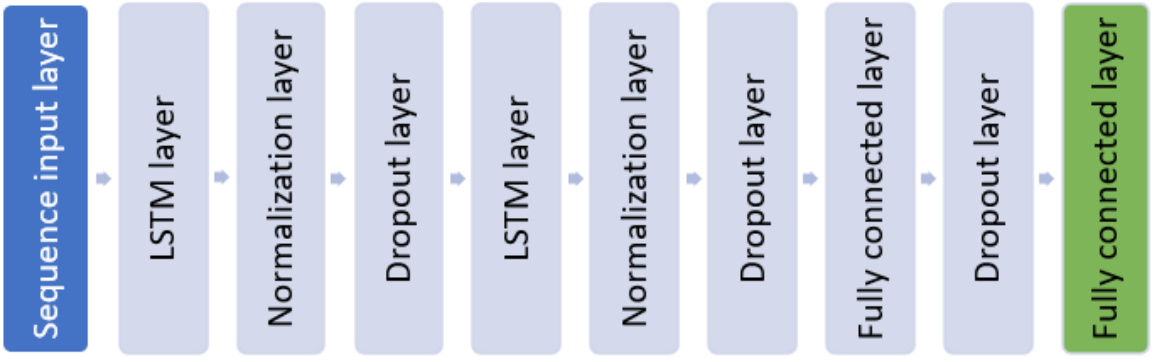


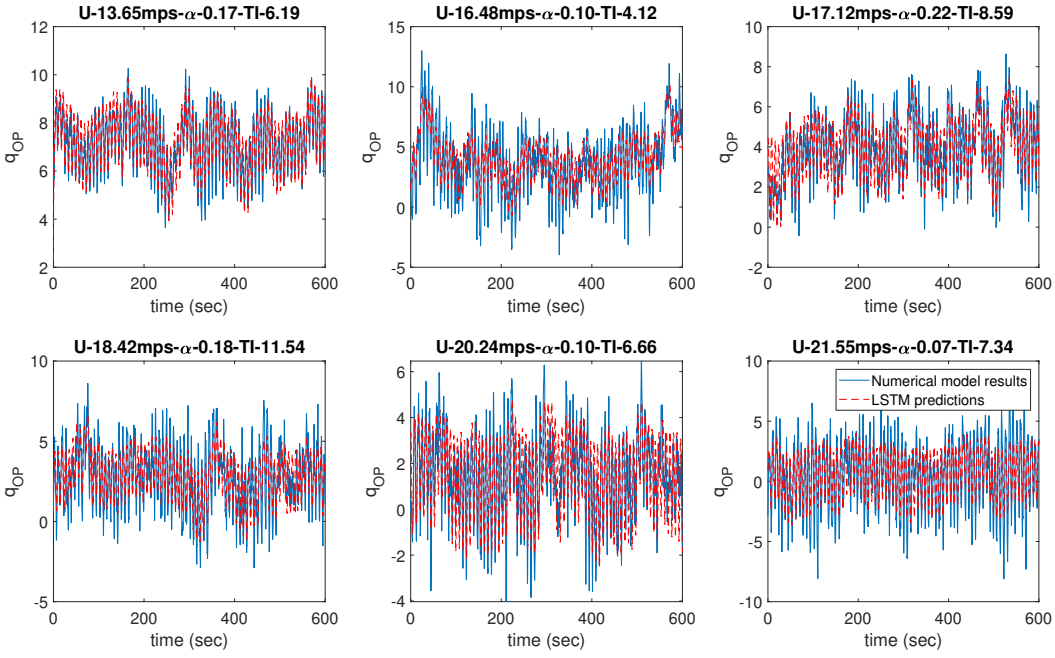
Figure 13. LSTM Model architecture for predicting blade response

constant features, principal components of the wind speed data were selected using the RFA algorithm. This approach enables accurate predictions of both in-plane and out-of-plane blade deformations (Fig. 14). Notably, the second principal component, when paired with angular blade position and blade pitch angle, yielded the best performance. This suggests that the second principal component likely captures patterns more relevant to predicting blade deformations than the first principal component, despite the first component capturing the dataset's maximum variance. These findings further reinforce the importance of using feature selection approaches in obtaining a simplified model representation with a small set of the most informative features. This method is further extended to predict the response at individual blade DOFs, which are the most fundamental quantities in numerical modelling, and the corresponding results are presented in 15. Using the fifth principal component as the input feature paired with blade angular position and blade pitch angle produced the lowest RMSE for predicting the response at individual blade DOFs. Here, it can be seen that a higher level of accuracy was obtained in the prediction of DOFs q_{B1F1} and q_{B1E1} as compared to q_{B1F2}. This can be attributed to the higher frequency of the mode shape for q_{B1F2} . Further, the absolute magnitude of q_{B1F2} shows that this DOF has lesser impact on the total deformation as compared to the response of q_{B1F1}. However, it is evident from Fig. 15 that a good approximation in predicting all the DOFs is achieved by capturing the governing dynamics in all the parameters.

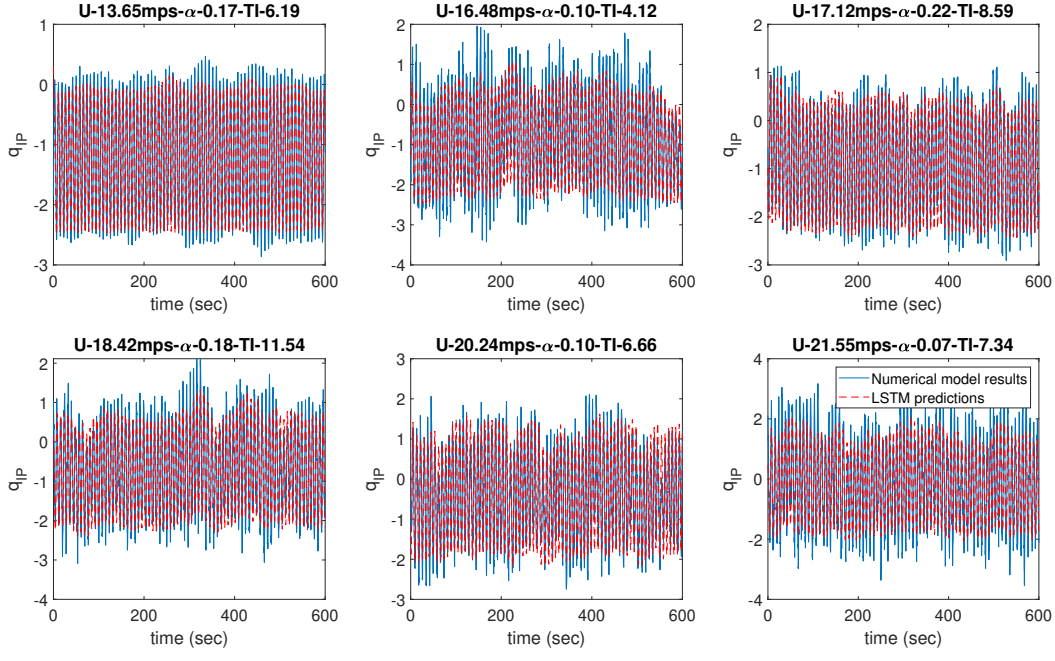
8.2.2 Blade response prediction with DCT features

Building upon the PCA approach, the use of DCT features for blade response prediction is explored in this section. The LSTM model follows the same model architecture presented in the previous section (Fig. 13). Following the methodology adopted in Section 8.2.1, the model architecture was trained using DCT features. Fig. 16 and Fig. 17 demonstrate the accuracy of DCT in predicting the blade response, exhibiting a close match to the true response. The best performance for blade response

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024



(a) Blade out of plane deflection response

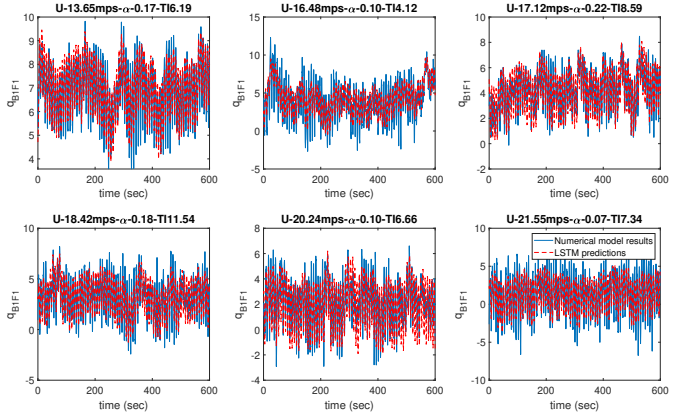


(b) Blade in-plane deflection response

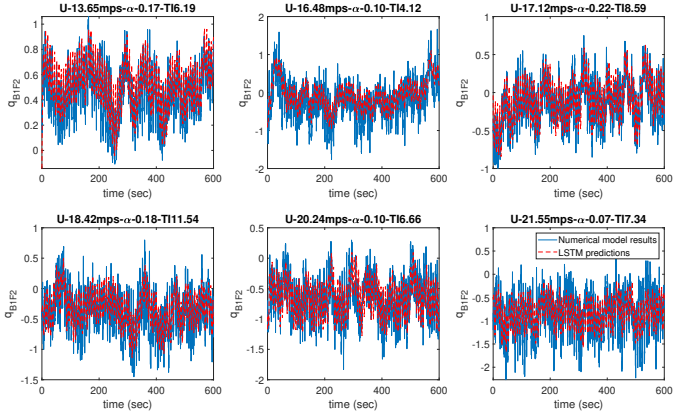
Figure 14. Multi-task learning LSTM Model predictions of blade deflections using PCA features

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

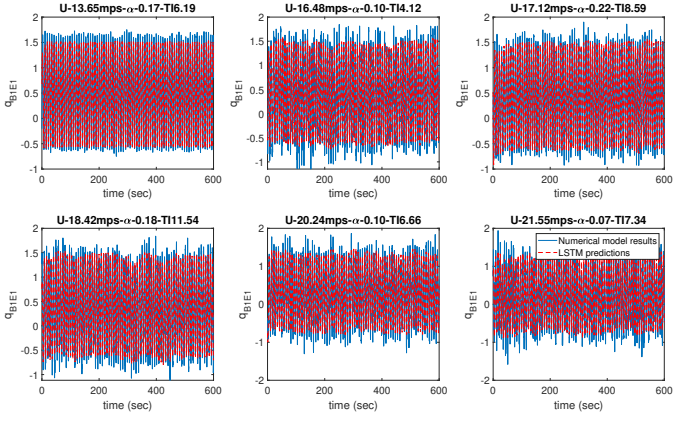
eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS



(a) Blade first flapwise DOF response



(b) Blade second flapwise DOF response



(c) Blade first edgewise DOF response

Figure 15. Multi-task learning LSTM Model predictions of blade DOFs using PCA features

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Output Parameter	RMSE	
	PCA	DCT
Blade IP Deformation	0.518	0.315
Blade Oop Deformation	0.595	0.378
Blade DOF -q_{B1F1}	0.602	0.565
Blade DOF -q_{B1F2}	0.665	0.637
Blade DOF -q_{B1E1}	0.469	0.454

Table 5. Comparison of LSTM Model performance trained using PCA and DCT features

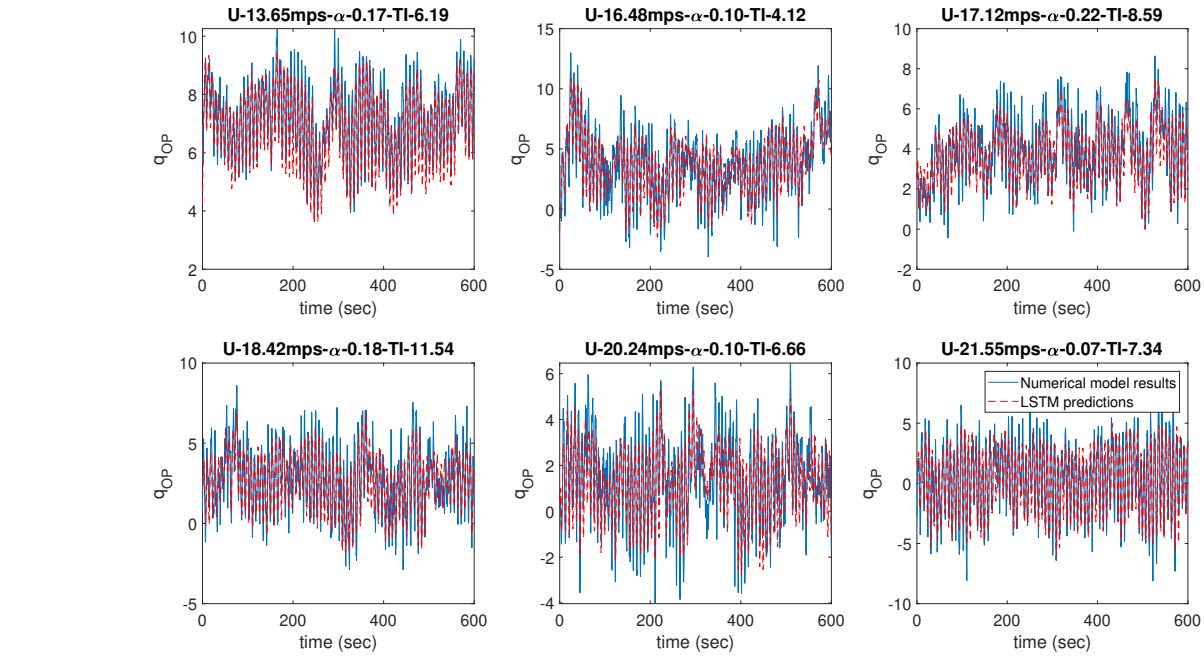
prediction was obtained using the second group of DCT components presented in Fig. 10 as input features, while the blade DOF predictions achieve higher accuracy under the combination of the second and third groups of DCT components.

Table 5 provides a comparison of RMSE obtained for models developed using PCA and DCT features, highlighting DCT's superior performance in predicting the blade response. These results highlights that LSTM model can accurately predict the dynamic response of wind turbine blades. The accuracy of predictions also depends on the dimensionality reduction algorithm and the features chosen for the model development. The accuracy achieved in the intermediate stage of the multi-stage modelling approach also significantly impacts the final model predictions. These results also demonstrate that DCT, by focusing on spatial frequency patterns, delivers better results in predicting the blade responses. The next section focuses on analyzing the computational advantage of the developed surrogate models.

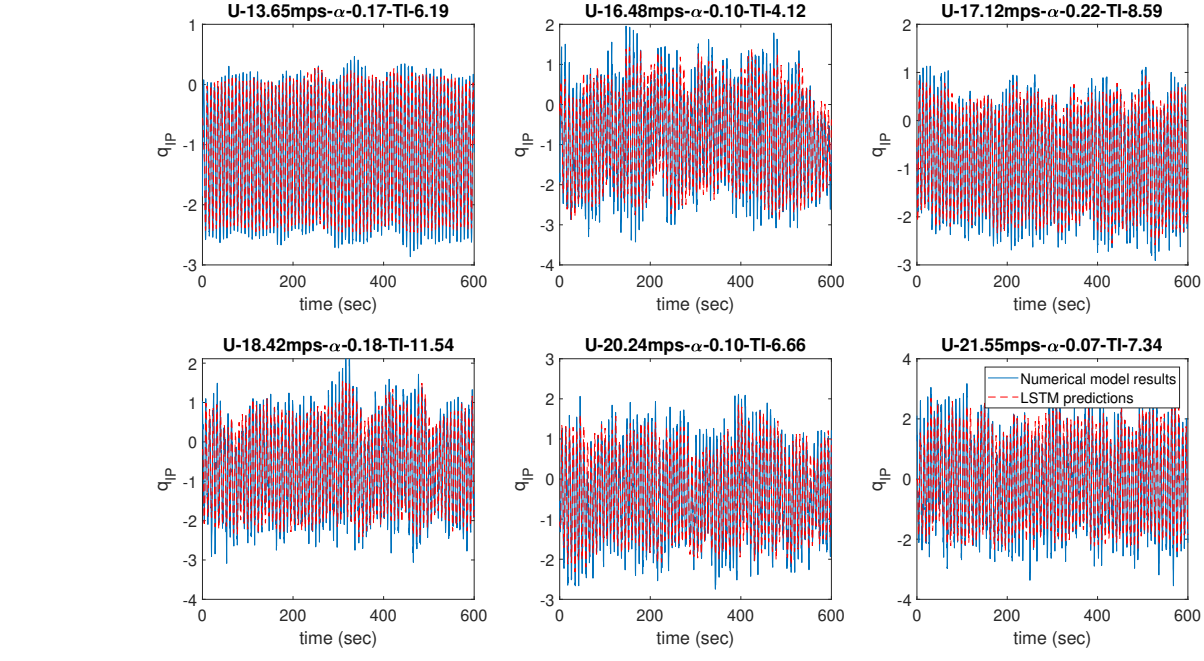
8.3 Computational advantage of surrogate approach

The need to reduce the computational cost in predicting the dynamic response of a wind turbine is at the heart of the surrogate model developed in this study. Having established the accuracy of the developed surrogate model, this section focuses on the computational advantages of using these models. A direct comparison of execution times was conducted to quantify the extent of the computational efficiency of these models. Both the numerical model and the surrogate model were used to predict blade response for a 600-second turbulent wind inflow generated using TurbSim. To ensure a fair comparison, only the essential blade DOFs and generator azimuth angle were activated within the MBD model. The surrogate modelling approach developed in this study offers a significant computational advantage in predicting blade response. The execution time comparison shows a 75-fold speed improvement using the surrogate model compared to the corresponding MBD model. This comparison highlights the computational efficiency of the developed surrogate models. The computational advantage allows for conducting a large number of simulations required for site-specific performance analysis. Furthermore, the ability of the surrogate model to quickly predict blade deformations and loads under varying wind and control inputs makes it a useful tool for comprehensive fatigue analysis of wind turbine blades. While quantifying the computational gains, all the simulations were performed on a system with an 8-core Intel Xeon CPU with a clock speed of 3.8GHz using 32GB RAM and running on Microsoft Windows 10 Pro.

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024



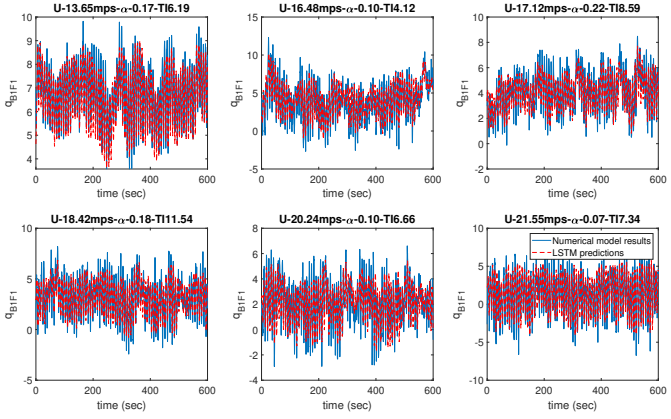
(a) Blade out of plane deflection response



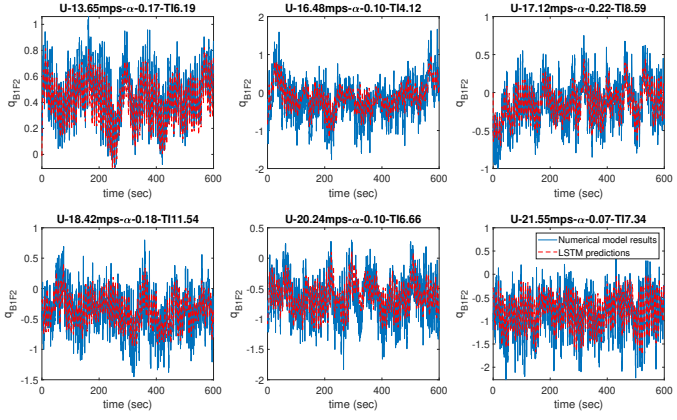
(b) Blade in-plane deflection response

Figure 16. Multi-task learning LSTM Model prediction of blade deflections using DCT features

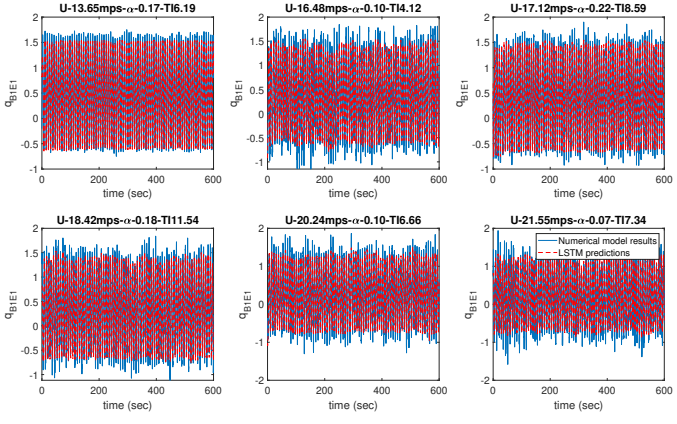
https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024



(a) Blade first flapwise DOF response



(b) Blade second flapwise DOF response



(c) Blade first edgewise DOF response

Figure 17. Multi-task learning LSTM Model predictions of blade DOFs using DCT features

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

9 Conclusions

In this manuscript, an approach that combines multi-stage modelling and multi-task learning with dimensionality reduction techniques and a feature selection algorithm was presented. This combined method aims to enhance the efficiency of developing LSTM models for predicting blade response. Based on the investigation performed in this paper, following conclusions can be drawn:

- Dimensionality Reduction: The effectiveness of PCA and DCT in simplifying wind field data while retaining crucial information for prediction tasks was demonstrated. Both PCA and DCT, particularly when combined with recursive feature addition, helped in achieving efficient model configuration and improved prediction accuracy.
- Rotor Speed and Blade Pitch Prediction: LSTM models were developed for rotor speed prediction (using PCA features) and blade pitch prediction (using DCT features). These models achieved varying degrees of accuracy in capturing the dynamics of their respective target parameters across different uncertainty levels in wind conditions. Accurate rotor speed and blade pitch information were identified as critical parameters for subsequent blade response prediction.
- Blade Response Prediction: A multi-stage modelling approach was employed, where predicted control parameters were fed into an LSTM model to forecast blade deformations and DOF responses. Models trained using DCT features showed higher accuracy as compared to PCA features for this task, indicating DCT's ability to capture spatial frequency patterns driving blade dynamics.
- Practical applications: The LSTM model presented in this paper are trained using input-output data only. This approach has potential applications in design feasibility studies for those models where the exact model configuration is not available due to the intellectual property concerns. Further, due to the low computational cost, these models can be used within model predictive control framework for regulating the performance.

Code and data availability. The code and 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.

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

Competing interests. There are no competing interests in the work presented.

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

600 *Acknowledgements.* This work is supported by Science Foundation Ireland (SFI) grant no. 20/FFP-P/8702.

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

References

Abbas, N., Zalkind, D., Pao, L., and Wright, A.: A reference open-source controller for fixed and floating offshore wind turbines, Wind Energy Science Discussions, 2021, 1–33, 2021.

Abbas, N. J., Zalkind, D. S., Pao, L., and Wright, A.: A reference open-source controller for fixed and floating offshore wind turbines, Wind Energy Science, 7, 53–73, 2022.

Bai, H., Shi, L., Aoues, Y., Huang, C., and Lemosse, D.: Estimation of probability distribution of long-term fatigue damage on wind turbine tower using residual neural network, Mechanical Systems and Signal Processing, 190, 110 101, 2023.

Baisthakur, S.: Multi-task Learning Long Short-term Memory Model to Emulate Wind Turbine Blade Dynamics, in: Wind Energy Science, Zenodo, https://doi.org/10.5281/zenodo.13305715, 2024.

Baisthakur, S. and Fitzgerald, B.: Physics-Informed Neural Network surrogate model for bypassing Blade Element Momentum theory in wind turbine aerodynamic load estimation, Renewable Energy, p. 120122, 2024a.

Baisthakur, S. and Fitzgerald, B.: Meta-modelling of wind turbine blades using machine learning based multivariate time series sequence modelling, submitted to Wind Energy, 2024b.

Banik, A., Behera, C., Sarathkumar, T. V., and Goswami, A. K.: Uncertain wind power forecasting using LSTM-based prediction interval, IET Renewable Power Generation, 14, 2657–2667, 2020.

Bashirzadeh Tabrizi, A., Wu, B., Whale, J., and Shahabi Lotfabadi, M.: Using TurbSim stochastic simulator to improve accuracy of computational modelling of wind in the built environment, Wind Engineering, 43, 147–161, 2019.

Calazone, O.: An Intuitive Explanation of LSTM, https://medium.com/@ottaviocalzone/an-intuitive-explanation-of-lstm-a035eb6ab42c, accessed on February 20, 2024, 2022.

Caruana, R.: Multitask learning, Machine learning, 28, 41–75, 1997.

Chen, H., Liu, H., Chu, X., Liu, Q., and Xue, D.: Anomaly detection and critical SCADA parameters identification for wind turbines based on LSTM-AE neural network, Renewable Energy, 172, 829–840, 2021.

Choe, D.-E., Kim, H.-C., and Kim, M.-H.: Sequence-based modeling of deep learning with LSTM and GRU networks for structural damage detection of floating offshore wind turbine blades, Renewable Energy, 174, 218–235, 2021.

de N Santos, F., D'Antuono, P., Robbelein, K., Noppe, N., Weijtjens, W., and Devriendt, C.: Long-term fatigue estimation on offshore wind turbines interface loads through loss function physics-guided learning of neural networks, Renewable Energy, 205, 461–474, 2023.

Dimitrov, N. and Göçmen, T.: Virtual sensors for wind turbines with machine learning-based time series models, Wind Energy, 25, 1626–1645, 2022.

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

Fitzgerald, B., McAuliffe, J., Baisthakur, S., and Sarkar, S.: Enhancing the reliability of floating offshore wind turbine towers subjected to misaligned wind-wave loading using tuned mass damper inerters (TMDIs), Renewable Energy, 211, 522–538, 2023.

Gaertner, E., Rinker, J., Sethuraman, L., Zahle, F., Anderson, B., Barter, G. E., Abbas, N. J., Meng, F., Bortolotti, P., Skrzypinski, W., et al.: IEA wind TCP task 37: definition of the IEA 15-megawatt offshore reference wind turbine, Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2020.

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

Garcke, J., Iza-Teran, R., Marks, M., Pathare, M., Schollbach, D., and Stettner, M.: Dimensionality reduction for the analysis of time series data from wind turbines, Scientific Computing and Algorithms in Industrial Simulations: Projects and Products of Fraunhofer SCAI, pp. 317–339, 2017.

Geng, D., Zhang, H., and Wu, H.: Short-term wind speed prediction based on principal component analysis and LSTM, Applied sciences, 10, 4416, 2020.

Haghi, R. and Crawford, C.: Surrogate models for the blade element momentum aerodynamic model using non-intrusive polynomial chaos expansions, Wind Energy Science Discussions, 2021, 1–20, 2021.

Haghi, R. and Crawford, C.: Data-driven surrogate model for wind turbine damage equivalent load, Wind Energy Science Discussions, 2023, 1–34, 2023.

Hamarneh, G., McIntosh, C., and Drew, M. S.: Perception-based visualization of manifold-valued medical images using distance-preserving dimensionality reduction, IEEE transactions on medical imaging, 30, 1314–1327, 2011.

Hochreiter, S. and Schmidhuber, J.: Long short-term memory, Neural computation, 9, 1735–1780, 1997.

Hou, C. K. J. and Behdinan, K.: Dimensionality reduction in surrogate modeling: A review of combined methods, Data Science and Engineering, 7, 402–427, 2022.

Hu, Y., Huber, A., Anumula, J., and Liu, S.-C.: Overcoming the vanishing gradient problem in plain recurrent networks, arXiv preprint arXiv:1801.06105, 2018.

Hübler, C., Gebhardt, C. G., and Rolfes, R.: Methodologies for fatigue assessment of offshore wind turbines considering scattering environmental conditions and the uncertainty due to finite sampling, Wind Energy, 21, 1092–1105, 2018.

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

Jonkman, B. J.: TurbSim user's guide: Version 1.50, Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2009.

Kane, T. R. and Levinson, D. A.: Dynamics, theory and applications, McGraw Hill, 1985.

Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980, 2014.

Lataniotis, C.: Data-driven uncertainty quantification for high-dimensional engineering problems, Ph.D. thesis, ETH Zurich, 2019.

Markaki, M. E. and Stylianou, Y.: Dimensionality reduction of modulation frequency features for speech discrimination., in: INTERSPEECH, pp. 646–649, 2008.

Njiri, J. G. and Söffker, D.: State-of-the-art in wind turbine control: Trends and challenges, Renewable and Sustainable Energy Reviews, 60, 377–393, 2016.

Pereira, T. P., Ekwaro-Osire, S., Dias, J. P., Ward, N. J., and Cunha Jr, A.: Uncertainty quantification of wind turbine wakes under random wind conditions, in: ASME International Mechanical Engineering Congress and Exposition, vol. 83501, p. V013T13A022, American Society of Mechanical Engineers, 2019.

Roga, S., Bardhan, S., Kumar, Y., and Dubey, S. K.: Recent technology and challenges of wind energy generation: A review, Sustainable Energy Technologies and Assessments, 52, 102 239, 2022.

Sarkar, S. and Fitzgerald, B.: Vibration control of spar-type floating offshore wind turbine towers using a tuned mass-damper-inerter, Structural Control and Health Monitoring, 27, e2471, 2020.

Sarkar, S. and Fitzgerald, B.: Use of kane's method for multi-body dynamic modelling and control of spar-type floating offshore wind turbines, Energies, 14, 6635, 2021.

Sarkar, S. and Fitzgerald, B.: Fluid inerter for optimal vibration control of floating offshore wind turbine towers, Engineering Structures, 266, 114 558, 2022.

https://doi.org/10.5194/wes-2024-105
Preprint. Discussion started: 14 October 2024

Sarkar, S., Fitzgerald, B., and Basu, B.: Individual blade pitch control of floating offshore wind turbines for load mitigation and power regulation, IEEE Transactions on Control Systems Technology, 29, 305–315, 2020.

Schär, S., Marelli, S., and Sudret, B.: Emulating the dynamics of complex systems using autoregressive models on manifolds (mNARX), Mechanical Systems and Signal Processing, 208, 110 956, 2024.

Shi, W., Hu, L., Lin, Z., Zhang, L., Wu, J., and Chai, W.: Short-term motion prediction of floating offshore wind turbine based on muti-input LSTM neural network, Ocean Engineering, 280, 114 558, 2023.

Skittides, C. and Früh, W.-G.: Wind forecasting using principal component analysis, Renewable Energy, 69, 365–374, 2014.

Sun, X., Huang, D., and Wu, G.: The current state of offshore wind energy technology development, Energy, 41, 298–312, 2012.

Tan, J. D., Chang, C. C. W., Bhuiyan, M. A. S., Nisa'Minhad, K., and Ali, K.: Advancements of wind energy conversion systems for low-wind urban environments: A review, Energy Reports, 8, 3406–3414, 2022.

Thrun, S. and Mitchell, T. M.: Learning one more thing, in: IJCAI, vol. 95, pp. 1217–1223, 1995.

Van Der Maaten, L., Postma, E. O., van den Herik, H. J., et al.: Dimensionality reduction: A comparative review, Journal of Machine Learning Research, 10, 13, 2009.

Wang, Y., Ma, X., and Joyce, M. J.: Reducing sensor complexity for monitoring wind turbine performance using principal component analysis, Renewable energy, 97, 444–456, 2016.

Woo, S., Park, J., and Park, J.: Predicting wind turbine power and load outputs by multi-task convolutional LSTM model, in: 2018 IEEE Power & Energy Society General Meeting (PESGM), pp. 1–5, IEEE, 2018.

Xiang, L., Wang, P., Yang, X., Hu, A., and Su, H.: Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism, Measurement, 175, 109 094, 2021.

Yu, R., Gao, J., Yu, M., Lu, W., Xu, T., Zhao, M., Zhang, J., Zhang, R., and Zhang, Z.: LSTM-EFG for wind power forecasting based on sequential correlation features, Future Generation Computer Systems, 93, 33–42, 2019.

Zhang, K., Tang, B., Deng, L., and Yu, X.: Fault detection of wind turbines by subspace reconstruction-based robust kernel principal component analysis, IEEE Transactions on Instrumentation and Measurement, 70, 1–11, 2021.

Zhang, Y. and Yang, Q.: An overview of multi-task learning, National Science Review, 5, 30–43, 2018.

Zhu, D., Huang, X., Ding, Z., and Zhang, W.: Estimation of wind turbine responses with attention-based neural network incorporating environmental uncertainties, Reliability Engineering & System Safety, 241, 109 616, 2024.