



Long-term fatigue estimation on offshore wind turbines interface loads through loss function physics-guided learning of neural networks

Francisco de N Santos^{a,*}, Pietro D'Antuono^a, Koen Robbelein^a, Nymfa Noppe^b, Wout Weijtjens^a, Christof Devriendt^a

^a OWI-Lab, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

^b 24SEA, Drukperstraat 4, 1000 Brussels, Belgium

ARTICLE INFO

Keywords:

Offshore wind
Structural health monitoring
Physics-informed machine learning

ABSTRACT

Offshore wind turbines are exposed during their serviceable lifetime to a wide range of loads from aero-, hydro- and structural dynamics. This complex loading scenario will have an impact on the lifetime of the asset, with fatigue remaining the key structural design driver for the substructure, e.g. the monopile. The ability to monitor the progression of fatigue life of these assets has recently become an operational concern. To achieve a monitoring alternative to strain gauges (cost-prohibitive farm-wide installation), supervisory control and data acquisition (SCADA) systems, often coupled with acceleration measurements, have been used. Existing work focused primarily on ten-minute fatigue load estimation. However, fatigue accumulates over time and the ability to accurately monitor this accumulation of fatigue over a longer time-window is paramount. In this contribution we investigate a novel approach using nine months of real-world SCADA and acceleration ten-minute statistics as inputs of a neural network model for long-term DEM estimation. This is further enhanced by including physical information relative to the problem at hand into the neural network model, in a so-called physics-informed machine learning approach. Specifically, we employ a custom loss function – the Minkowski logarithmic error – which prioritizes conservativeness (over-prediction of fatigue rates) and to embed the damage accumulation into the machine learning model.

In the results and discussion section, we use Tower-TP interface load measurements for three real-world turbines to demonstrate the concept. First a model is trained on one turbine, before being applied to all three locations. The model performance is compared to direct measurements of fatigue progression at all three locations and a control loss function (the mean squared logarithm error) in both the Fore-Aft and Side-Side loading directions. The novel physics-informed model clearly outperforms the control loss function in accumulated fatigue predictions with errors below 3% on accumulated damage equivalent moment estimates. Additionally, the model's long-term performance is checked according to varying timescales (hourly, daily, weekly or monthly accumulation of DEM) and it is seen that the spread on the error rapidly reduces and converges to zero. Finally, the long-term DEM is accumulated into a monthly accumulated fatigue damage. We also see how the error on this begins to converge after 6+ months, which opens the door for extrapolation.

1. Introduction

1.1. Motivation

Recent years have seen a surge in the demand for wind energy as societies attempt to transition to sustainable energy sources. To meet the promised lowered costs, more cost-effective designs are being targeted. These cost-reductions are in part achieved through the increased average wind turbine size (from 4.8 MW in 2016 to 11.2 MW of average power rating of ordered offshore turbines by 2021 [1,2]). In parallel, recent insights have identified (unintended) conservatism

in older offshore wind farms, and updated design procedures (e.g. PISA [3]) have allowed for more accurate predictions of structural behaviour.

While offshore wind turbine design has seen much evolution, fatigue has remained a key design driver [4]. During design substructure dimensions are optimized for fatigue life to match the project's intended lifetime (typically 20–25 years) as close as possible. However, in older designs, in particular those that are now reaching their intended half-life, a structural reserve has been found. Real-world observations through the use of Structural Health Monitoring (SHM) seem to

* Corresponding author.

E-mail address: francisco.de.nolasco.santos@vub.be (F. de N Santos).

<https://doi.org/10.1016/j.renene.2023.01.093>

Received 31 August 2022; Received in revised form 29 December 2022; Accepted 25 January 2023

Available online 28 January 2023

0960-1481/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Abbreviations and Nomenclature

Abbreviations

e.g.	<i>Exempli gratia</i> (for example)
i.e.	<i>Ied est</i> (that is)
ANN	Artificial Neural Network
API	Application programming interface
CI	Confidence interval
CoE	Cost of Energy
DEM	Damage Equivalent Moment
EOP	Environmental and Operational Parameters
FA	Fore-aft
GELU	Gaussian Error Linear Unit
KDE	Kernel density estimator
LAT	Lowest Astronomical Tide
ML	Machine Learning
MLE	Minkowski Logarithmic Error
MSLE	Mean Squared Logarithmic Error
O&M	Operation and Maintenance
OWT	Offshore Wind Turbine
PDF	Probability density function
PINN	Physics-Informed Neural Networks
PSD	Power spectral density
ReLU	Rectified Linear Unit
RPM	Rotations Per Minute
RUL	Remaining Useful Lifetime
SCADA	Supervisory Control and Data Acquisition
SHM	Structural health monitoring
SN	Stress-cycle curve
SS	Side-to-side
T	Training Turbine
TW-TP	Tower-Transition Piece interface
XL	Extra Large

Nomenclature

\hat{Y}	Predicted values vector
Y	Measured values vector
δ_D	Long-term fatigue damage error
δ_{LT}	Long-term DEM error
log	Logarithmic function
\mathbb{R}^n	n -dimensional real vector space
D	Dropout rate set
\mathcal{L}	Loss function
\mathcal{O}	Optimizer learning rate set
ω	Rotational speed of rotor (RPM)
\bar{r}_i	Mean residuals
Φ	Physics-informed
σ	Stress range, standard deviation
θ	Yaw angle (°)
$\hat{p}_n(y)$	Probability density function
DEM_{LT}	Long-term DEM
L	Loss function
L^p	Lebesgue space norm of p -dimension

m	Wöhler exponent
M_{fl}	Lateral bending moment
M_{fn}	Normal bending moment
n	Number of cycles
N_{eq}	Equivalent number of cycles (predefined)
p	Lebesgue space dimension
$P_n(y)$	Cumulative distribution function
R^2	Coefficient of determination
r_i	Residual
r_i	TW/TP inner radius at strain gauge location
r_o	TW/TP outer radius at strain gauge location
x	Real number vector

that the initial capital expenditure required for SHM is quickly offset by the reduction in operational expenditures [7], complemented with the possibility of lifetime extensions or optimized usage.

In more recent projects this structural reserve has seemingly been diminished as the design codes were optimized. As a result, operators want to keep tabs on the fatigue progression of their assets to assure intended project life can be reached. This, in particular when external factors interfere with the normal operation of the assets, e.g. shutting down a turbine to meet demands from the transmission system operator, but that might be detrimental to expected fatigue life. In absence of any reserve, to monitor fatigue life has now become an operational concern.

It should, however, be noted that fatigue monitoring is not a substitute, but rather a complement, of fault diagnosis/prognosis [8,9], wherein the structural fatigue health monitoring allows operators to make long-term decisions regarding their assets' predicted life and fault prognosis aids operators identifying problematic components and plan maintenance ahead of time.

1.2. State-of-the-art in fatigue monitoring

The accurate monitoring of the remaining useful lifetime (RUL) for every turbine within a given wind farm can thus help wind farm operators of both older and newer farms to take informed decisions regarding the operation of their assets.

Historically, this assessment has been performed by monitoring the assets, either through physical inspections [10] or through strain gauges, part of an SHM system, installed at the substructures' interface with the turbine, capable of capturing the turbines' fatigue load history [11–13]. However, farm-wide physical inspections are costly and workplace-accident prone. As for strain gauges, these remain labour intensive to install and prone to failure within a wind turbine's operational life. Installing and maintaining strain gauges on the substructure of every wind turbine is considered cost-ineffective. As a consequence, real-world examples known to the authors have less than 10% of turbines in a farm equipped with a SHM system which includes strain gauges.

Several researchers have therefore previously suggested to rely instead on the data obtained from the supervisory control and data acquisition (SCADA) system installed in every turbine [14–16]. Their strategy typically comprises in training models to predict fatigue loads as measured on turbines with a SHM system from SCADA data. As SCADA is available on every turbine, one can use said models to estimate fatigue life on every location [17,18]. [18] performs a comparative study between different sensor setups for fatigue estimation, considering varying quality of SCADA and accelerations, results show a strong improvement in the ability to estimate fatigue rates when tower acceleration data was considered. This result may come to no surprise as both tower accelerations and fatigue cycles follow the same

suggest that fatigue life consumption is less than those computed during design [5]. Motivated by a quest for greater cost-effectiveness, this observation can optimize operation and maintenance (O&M), which amounts to nearly a third of global costs [6]. It has been demonstrated

equations of motion of the OWT [19]. Currently, the cost of farm-wide deployment of high quality accelerometers has dropped far below that of a farm-wide deployment of strain gauges or inspection-based monitoring. Moreover, as accelerometers are a reliable technology, it does not require periodic inspections carried out by technicians (reducing workplace accident probability).

In parallel, an increased use of machine learning (ML) algorithms and other data-driven techniques has been noted, producing promising results. These have included proofs of the reliability of data-driven fatigue estimators serving the entirety of the turbine's operational life [20], the quality of SCADA-based artificial neural network (ANN) models in different flow conditions for blade flap- and edgewise bending moment fatigue load estimation [21,22], the use of ANNs for blade root flapwise fatigue loads and their connection to turbine failure [23], Gaussian process regression time-series modelling to evaluate the influence so-called EOPs (Environmental and Operational Parameters) have on the features of the vibration response of the wind turbine blades [24], the use of conditional variational auto-encoders to estimate the probability distribution of the accumulated fatigue on the root cross-section of a simulated wind turbine blade, enabling long-term probabilistic deterioration predictions based on historic SCADA data [25,26] or of graph neural networks [27], the use of long-term SCADA from onshore wind turbines to estimate the tower fore-aft bending moment [28], the use of Gaussian processes for damage detection [29] and the use of SCADA and acceleration data to predict the fatigue loads at tower-transition piece on a ten-minute level and provide insight into feature selection, the performance of different sensor setups and a tentative farm-wide employment [18].

A caveat of the aforementioned research is that predictive quality of the models is assessed on typically a ten-minute basis. *I.e.* a model quality is evaluated by its ability to predict ten-minute damage rates. However, in offshore wind, the evaluation of fatigue life is not primarily performed on a ten-minute basis, but rather by assessing the total accumulation of the fatigue damages over the project life and quantifying residual life [5,30]. The present contribution will show that the ability to predict ten-minute fatigue rates well, does not guarantee neither an accurate nor a conservative outcome when these predictions are accumulated into a fatigue life estimation. To resolve this caveat the present contribution will introduce physics-informed machine learning models (Φ -ML) [31] to significantly improve the accuracy of fleet-wide fatigue life predictions by incorporating damage accumulation into the ML model itself. Simultaneously, conservatism is introduced into the predictions by using the properties of the logarithmic function.

For this objective, data collected from XL monopile foundations is utilized; this represents a radical departure in terms of the structure's dynamics when compared to smaller monopiles or jackets, with an increased importance of side-to-side loading and loading under idling conditions [32,33]. As such, an added focus will be given to the performance of the proposed methodology's in diverse operational conditions.

1.3. Physics-informed machine learning

The recent increase in monitoring data availability has led to a predominance of data-driven approaches in SHM, the inability for these approaches to properly extrapolate or to explicitly take into consideration specific physical relationships of the underlying problem limit their applicability. This has led to the rise of physics-guided machine learning (Φ -ML) approaches.

In Φ -ML, hybrid models are generated which are not completely data-driven and include some underlying physical knowledge of the problem at hand. As good as purely data-driven models may fit the training data, some predictions may be physically nonsensical or inconsistent. This can be due to extrapolation (for which ML is not usually adequate) or observational biases that may lead to poor generalization [31]. Therefore, a drive for fundamental physical laws integration

and domain knowledge has recently risen, wherein ML models are taught the governing physical laws which provide 'informative priors' in the form of strong theoretical constraints and inductive biases on top of the observational ones [31].

More generally, one can understand Φ -ML as corresponding to a series of methods halfway in the spectrum between 'white-box models' (purely physics-based models, as partial differential equations or finite element methods) and 'black-box models' (where the model structure is purely data-driven). Therefore, several authors have spoken of 'grey-box' or hybrid models [34].

As mentioned, there is a spectrum of approaches between physics-principled models and data-driven models, from residual models and hybrid architectures to physics-guided learning. A quick subdivision of these diverse approaches can be performed by grouping these by the type of bias they incorporate: Observational biases, where reliable, high-quality data that embodies the underlying physics (through sensors) and carefully crafting input augmentation procedures (such as metrics, statistics and nonlinear transformations); inductive biases, where prior assumptions are incorporated by tailored interventions to an ML model architecture, such that the predictions sought are guaranteed to implicitly satisfy a set of given physical laws, typically expressed in the form of certain mathematical constraints. Albeit the most principled Φ -ML implementation, it often leads to complex implementations of difficult scalability [31]; and learning biases – the focus of this contribution – which consist in appropriately choosing loss functions, constraints and inference algorithms that modulate the training phase to explicitly favour convergence towards solutions that adhere to the underlying physics. Therefore, a flexible soft constraint platform which approximately satisfies the underlying physics is introduced, incorporating physics-based biases, *e.g.* physics-informed neural networks (PINNs [35]), as will be discussed in this contribution.

This approach is particularly relevant for our application, as we are attempting to have similarly acceptable results for ten-minute predictions and the long-term accumulation of these predictions. Additionally, when we consider a fleet-leader model (*i.e.* when a model trained on one turbine is considered representative of the entire fleet and used for farm-wide estimation [17]), the successful transferability of a model trained on an individual turbine will depend on the similitude of the input variable space throughout the farm. The additional stability/robustness added by Φ -ML may allow to extend machine learning into extrapolation scenarios.

1.4. Article structure and main contributions

The current article presents as its main contributions the development of a data-driven methodology which is able to tackle different timescales for DEM estimation (10-minute and long-term accumulation) through physics-guided learning by implementing a custom loss function. The structure of the article is divided into five individual sections, with the current section consisting in the introduction. In Section 2 the data and fatigue calculation methodology are presented, firstly by focusing on the data acquisition systems (Section 2.1) and after by introducing the mathematical background behind damage equivalent loads and fatigue life estimation (Sections 2.2 and 2.3). Section 3 describes the methodology for long-term fatigue estimation using physics-informed neural networks, with a greater focus on the loss function (Section 3.1). Section 4 presents the results of this contribution along with its discussion. More specifically, Section 4.1 focuses on the ten-minute level performance, Section 4.2 focusing on the long-term performance and Sections 4.2.1 and 4.2.2 focus on identifying the sources of model shortcomings. Section 4.3 presents the performance for different operational conditions, Section 4.4 for different timescale accumulations and Section 4.5 show the long-term fatigue damage accumulation. Finally, in Section 5 the conclusions and future work are discussed.

Table 1
Description of datasets from the measurement campaign. Note the conflicting sampling frequencies. Each data-type is processed into 10-min target statistics.

	Sensor	Sampling frequency	Variable	Units	
Input	SCADA	1 Hz	Rotational speed	rpm	
			Yaw angle	deg	
			Pitch angle	deg	
			Power	kW	
			Wind speed	$m\ s^{-1}$	
		Wind direction	deg		
		Target statistics (10-min): mean, minimum, maximum, standard deviation.			
	Wave buoy	10-min	Wave height	cm	
			Average wave period	s	
			High frequent wave direction	deg	
Low frequent wave direction			deg		
Tidal level			cm		
10% highest waves			cm		
	Height waves with period > 10 s	cm			
	Target statistics (10-min): mean.				
Accelerometers	$\left\{ \begin{array}{l} 16\text{ m LAT} \\ 69\text{ m LAT} \\ 97\text{ m LAT} \end{array} \right\}$	12.5 Hz	FA acceleration	g	
			SS acceleration	g	
			X (225°) acceleration	g	
			Y (225°) acceleration	g	
	Target statistics (10-min): mean, minimum, maximum, root mean square.				
Target	Strain gauges	30 Hz	Normal bending moment (M_{in})	MNm	
			Lateral bending moment (M_{il})	MNm	
	Target statistics (10-min): damage equivalent moments (DEM_{in} and DEM_{il}).				

2. Sensors, data and methodology

2.1. Sensors and data

In the present contribution, data from a measurement campaign on real-world offshore wind turbines with XL monopile foundations (9.5 MW turbines and water depths of up to 36 m) located in the Belgian North Sea was used. This is highly relevant, as the vast majority of offshore wind turbine foundations are monopiles (up to 80% [36]) and the increase in size of wind turbines, enabled by improvements in manufacturing processes, has led to the existence of very large monopiles — XL monopiles. A lower threshold of 8 m diameter (capable of supporting a 7–8 MW turbines) can be set for this term [37].

The growing size of turbines has led to a novel paradigm for the industry, as the natural frequencies of turbines with XL monopile foundations are getting increasingly closer to the wave frequency spectrum [38]. This, combined with a larger surface area for hydrodynamic loads and deeper locations, has led to a greater impact of the wave loading on fatigue [32]. This in turn has resulted in the increased importance of idling operational conditions and the inclusion of side-to-side (SS), or cross-wind, damage calculations. As all of this can be linked to the structural dynamics of the turbines, acceleration data has become paramount to understand the broader fatigue picture [33].

The measurement campaign was based on data collected by three types of sensors: 6 axial strain gauges used to calculate the bending moments (M_{in} , M_{il}) and installed along the inner circumference at the interface between the turbine tower and the transition piece (TW-TP), three dedicated tower bi-axial accelerometers installed at three different levels in the bottom, middle and top of the turbine tower and SCADA data collected at nacelle-level. The yaw angle (from SCADA) can be used to convert the bending moments from the strain gauges into FA (M_{in}) and SS (M_{il}) directions, as these are the ‘primary’ directions of the structure, with dynamics clearly being different between these two primary directions. It should be noted that the data was collected by third-parties, namely the operator when it comes to SCADA and a specialized monitoring company for the monitoring data. Throughout the monitoring data collection, extensive calibration and quality checks were performed. Some periods had no collected data due to power

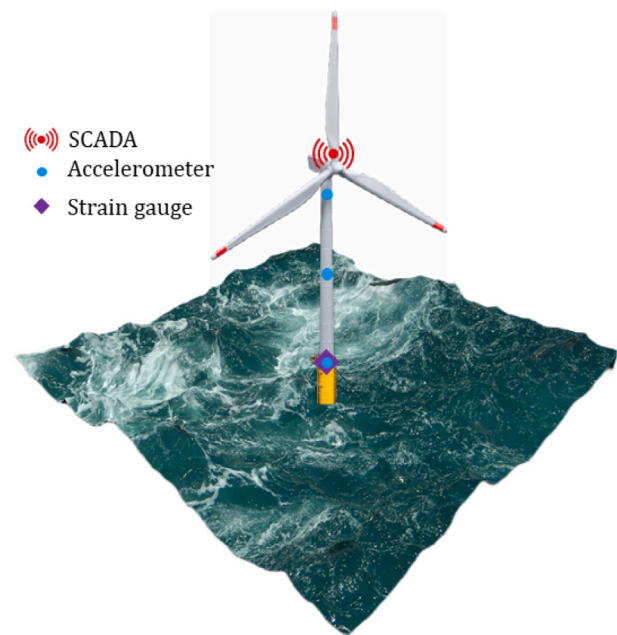


Fig. 1. Illustration of the sensor setup.

outage which does not have an impact in this contribution’s objectives; however, it should be considered when extrapolating in time. A schematic representation of the setup can be seen in Fig. 1.

The signals of the various sensors are collected at various sampling frequencies and transformed in ten-minute time instances of diverse statistics of the signals. A full description of the various parameters and their original sampling frequency can be found in Table 1. It should be noted that, albeit the sampling frequency for SCADA is 1 Hz, only ten-minute statistics are available (due to data storage considerations). [39] addressed the shortcomings a mean ten-minute SCADA approach might entail, however, the inclusion in this contribution of additional statistics such as minima, maxima and standard deviation

serves to rectify some of these by increasing the valuable information stored from SCADA signals.

Three turbines were instrumented, each representing one of the design clusters in the original detailed design of the wind farm. We can see the use of data from three turbines of different design clusters as a step in the direction of the farm-wide implementation of a fleet-leader model. The data collected in these locations was further enhanced by the addition of wave and tidal data, attained from a public Flemish maritime database (Meetnet Vlaamse Banken [40]), with a measuring station in the Belgian North Sea at the Westhinder site.

2.2. Damage equivalent moments

The strain gauges installed along the inner circumference at the TW-TP interface level enable the calculation of the bending moments in FA (M_{in}) and SS (M_{tl}) directions, through the ten-minute average yaw angle from the SCADA. Given the timeseries of bending moments, one can then employ a rainfall counting algorithm [41,42], which counts the number of cycles within a given stress range. Then, holding the linear damage accumulation hypothesis as true (Palmgren-Miner's rule) [43] and through the employment of the Wöhler exponent (the negative inverse slope of the SN curve) [44], the damage equivalent moments (DEM) are calculated for any ten minute window. A more detailed discussion of this procedure can be found in [45]. These DEMs will become the input for a fatigue assessment of the asset and are fairly widely adopted among load engineers to quantify fatigue rates (over e.g. damage). As DEMs are also widely used in design, results can be compared with design results (when available).

The equation for the DEM is given in Eq. (2.1), as defined by [46] (here presented for the stress ranges), wherein m is the inverse slope of the SN curve, or Wöhler exponent [42], n_i is the number of cycles of a given stress range, σ_i , r_o and r_i , are respectively the TW/TP outer and inner radii at strain gauge locations. For this contribution, and following the farm's design documentation, the value of 5 was used for m and $N_{eq} = 10^7$, a predefined number of cycles.

$$DEM = \frac{1}{N_{eq}} \cdot \left(\sum_i n_i \cdot \left(\frac{\Delta\sigma_i \cdot \frac{\pi}{2} \cdot (r_o^4 - r_i^4)}{r_o} \right)^m \right)^{1/m} \quad (2.1)$$

The DEMs are calculated for both the FA (DEM_{in}) and SS (DEM_{tl}) directions using the respective timeseries.

As explained above, the DEM calculation hinges upon the bending moments estimated from the strain gauges' readings. As these are not available farm-wide, our strategy is based on the use of not only SCADA, but also accelerations. Accelerometers have proven to be a far more reliable sensor compared to strain gauges over past SHM projects. Moreover, unlike strain gauges, their installation is fairly simple and post-processing is straight-forward, without the need for temperature compensation or calibration. The strong correlation between strain and accelerations (there is a link between the measured horizontal accelerations and the bending stresses [19], and as seen in Fig. 2) allows us to justify this approach, as the influence of the fatigue damage on the structure appears to also be captured by accelerometers. In Fig. 2(b) we can see how the first order dynamics dominate both the accelerometer's and the strain gauge's signals.

2.3. Fatigue life estimation

DEM's are a valuable property when discussing a lifetime assessment procedure. Nonetheless, just calculating these on a ten-minute basis is not enough to actually provide an insight into the fatigue life of an asset. For this, short-term ten-minute DEMs have to be added up.

As seen in the previous section, DEMs translate the damage caused by a dynamic load history into the single-amplitude equivalent load which would cause the same damage. By Palmgren-Miner's rule, we can further combine n equivalent load ranges that have been derived for the

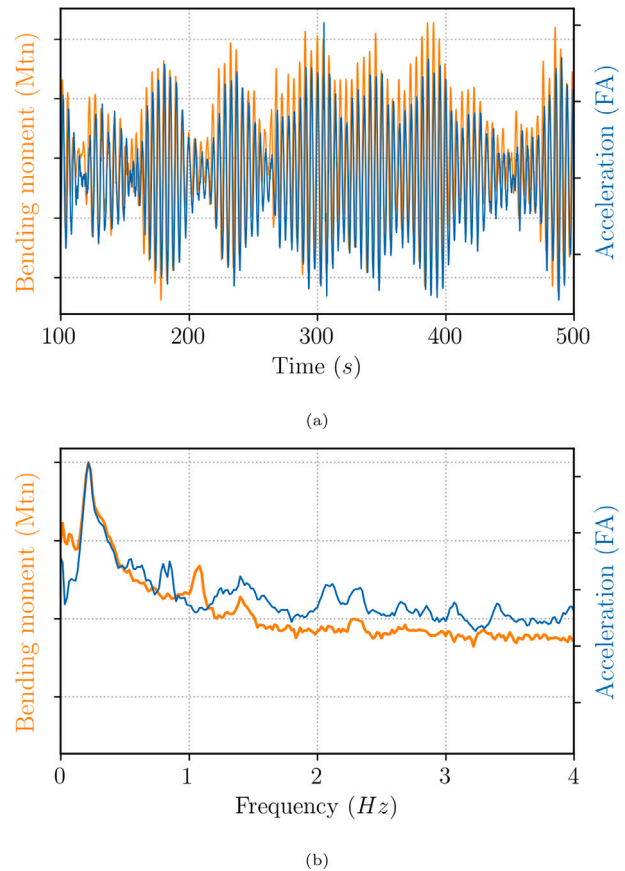


Fig. 2. Exemplary datasets of both the bending moment (orange) and the top level accelerations (blue) for a turbine in parked conditions. In (a), signals' timeseries, in (b) their corresponding power spectral densities (PSD). It can be clearly be observed that the two signals show a very similar behaviour, illustrating the strong link between the acceleration and fatigue data.

same reference cycle number and Wöhler exponent through the m -root of the weighted summation of the m -power DEM instance [47], as seen in Eq. (2.2). Here DEM_{LT} represents the long-term DEM accumulation of i ten-minute DEM instances. In this equation every i ten-minute time-instance DEM represents a damage load with identical occurrence probability of $1/n$ [48].

$$DEM_{LT} = \sqrt[m]{\frac{1}{n} \sum_{i=1}^n DEM_i^m} \quad (2.2)$$

As we can see, whatever the desired target long-term timescale might be, the combination of DEMs will always re-scale these into the timescale of DEM_i (all n instances of DEM_i must have the same timescale and have been calculated for the same Wöhler exponent, m). This allows to compare long-term effects of different timescales and easily add these by applying a timefactor. To translate the accumulated DEM (DEM_{LT}) into accumulated fatigue damage (D), one must simply re-instate Palmgren-Miner's rule and bring DEMs back into damage, as shown in Eq. (2.3).

$$D = \frac{DEM^m \cdot N_{eq}}{\bar{a}} = \frac{N_{eq}}{\bar{a}} \cdot \left(\frac{DEM \cdot r_o}{I} \right)^m, \quad (2.3)$$

where I is the area moment of inertia [13,49].

All of these quantities are known and have been introduced in previous sections apart from the intercept of the SN curve, \bar{a} . This value can easily be consulted in the corresponding SN-curve tables used during design [49]. In this case, we take the \bar{a} value from the SN curve D for seawater environment with cathodic protection, with $\bar{a} = 15.606$.

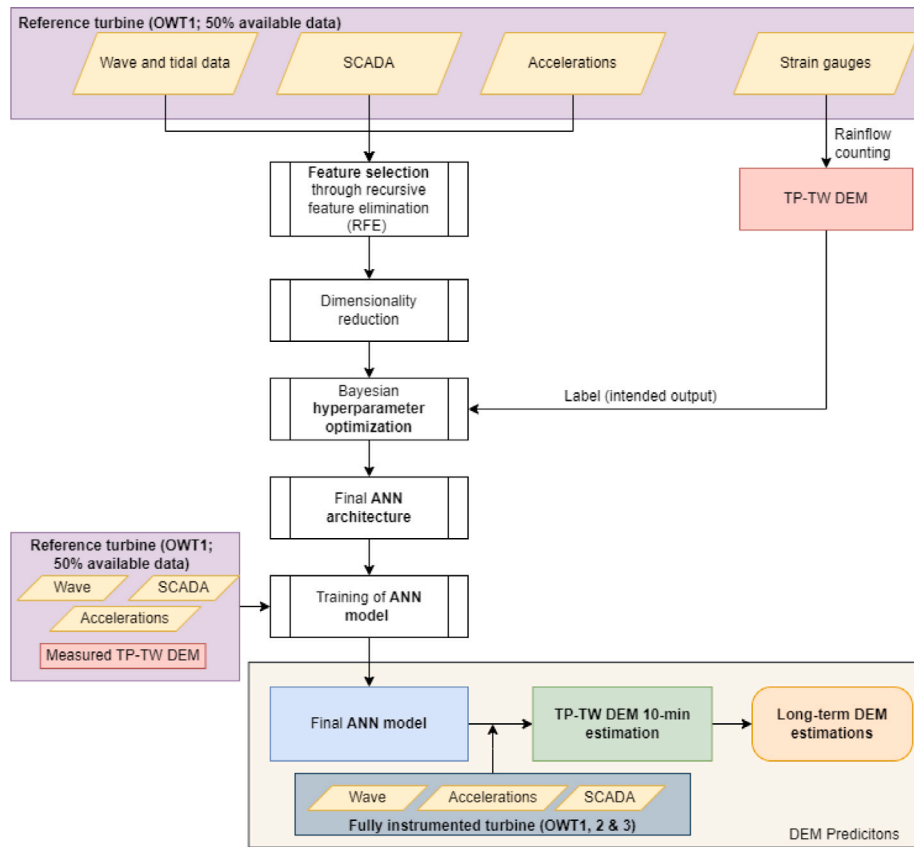


Fig. 3. Schematic of methodological approach.

Finally, we can translate the accumulated fatigue damage into the remaining useful lifetime through Eq. (2.4), where t_{op} is the number of years the wind turbine has already operated [13].

$$RUL = \frac{t_{op}}{D} - t_{op}, \quad (2.4)$$

3. Long-term fatigue estimation using physics-informed neural networks

In the present contribution, nine months worth of data from one real-world turbine (OWT1) is used to train ANNs in order to predict the DEMs affecting offshore wind turbine structures at TW-TP interface level for FA and SS directions. This mirrors the methodology prescribed in [18]. The full methodology can be seen in Fig. 3 and is applied to both FA (DEM_{tm}) and SS (DEM_{tt}) independently.

As can be seen in Fig. 3, variable-space dimensionality reduction on half of the available data from the reference (OWT1) was performed. [50] performed a systematic comparative study of different feature selection algorithms and explains the coalescing into the best performing feature selection algorithms, namely, Borutashap [51] and Recursive Feature Elimination [52], with the latter producing a smaller number of explanatory features, and thus being the selected algorithm.

The features selected for each of the two direction differed. Both FA and SS had accelerations highly represented, respectively relying on FA and SS accelerations. In the FA direction additional selected features were more related to the output of the turbine (such as wind speed, power, etc.), whereas SS fatigue loads are more defined by wave-related parameters (as high frequent wave direction, wave height and tidal level). A more detailed discussion on the use of feature selection can be consulted in previous works [18,53].

The general architecture of the machine learning model is that of a feed-forward artificial neural network (ANN). Using the `keras-tuner` package, three hyperparameter optimization algorithms (Random search [54], Hyperband [55] and Bayesian optimization [56]) are

tested, with the best model architecture being obtained with Bayesian optimization [57] with a mean squared error objective using the `keras-tuner` package [58]. The topology of the network consisted in 5 hidden layers, with the number of neurons varying between 32 and 512, rectified linear unit (ReLU) [59] and Gaussian error linear unit (GELU, a smoothing of ReLU using $\Phi(x)$, the cumulative distribution function of the Gaussian distribution [60]) activation functions. The dropout rate could be chosen from the set $D = \{0, 0.1, 0.2, 0.3\}$ and the learning rate of the optimizer (Adam) from $\mathcal{O} = \{1 \cdot e^{-2}, 1 \cdot e^{-3}, 1 \cdot e^{-4}\}$. The mean squared error and mean absolute error were the monitored metrics.

For training, half of the available data of the reference turbine is bootstrapped and the neural network model trained using a k-fold cross-validation [61] with 5 folds and a 70–30 validation split. A fixed threshold of 100 epochs is given and an early stop callback employed.

After training, the FA and SS ANNs were used to estimate the DEM both for the full dataset of the reference (training) turbine, OWT1, but also for two additional fully instrumented turbines (OWT2 and OWT3). Additionally, the sum of all the 10-min DEL estimations (and measurements) was scaled into a 10-min lifetime DEL (*cf.* Eq. (2.2), derived from [48]).

3.1. Loss functions

In Section 2.2 we have seen how damage equivalent loads on a ten-minute level can be accumulated and re-scaled back in order to compare fatigue damage throughout any timescale by a so-called long-term DEM (DEM_{LT}). While previous research focused on accuracy at a ten-minute level [18], the current contribution prioritizes instead the performance on long-term DEM. This, while attempting to retain accuracy on a ten-minute level and keeping a reasonable degree of conservativeness. Simultaneously an inherently conservative strategy is

desired, in which an overestimation of fatigue damage (and therefore an underestimation of RUL) is highly preferred over underestimating fatigue damage.

This multi-timescale (ten-minute predictions, long-term accumulation) is a highly complex problem statement, and a common ML approach will not be able to capture all the dynamic relationships between the many physical variables varying across space and time at different scales [62]. Therefore, in order to reflect the focus on long-term DEM estimation, a physics-informed machine learning approach is required, by guiding the learning process of the neural network, as introduced in Section 1.3.

In this section, two loss functions are selected and compared: the mean squared logarithmic error, a commonly available loss function which will serve as the control function, and a loss function, named *Minkowski logarithmic error*, custom-developed as to contain some physical information inherent to the problem at hand and reflect the priority given to long-term DEM estimation. Both functions employ the logarithmic function, which will inherently favour overpredictions, rendering the models conservative when used for a fatigue life assessment.

3.1.1. Mean squared logarithmic error

The mean squared logarithmic error (MSLE) can be interpreted as a measure of the ratio between the true and predicted values and employs the logarithmic function to the mean squared error loss function (see Eq. (3.1), where $\mathbf{Y}, \hat{\mathbf{Y}}$, with $\mathbf{Y} = (y_1, \dots, y_n)$ and $\hat{\mathbf{Y}} = (\hat{y}_1, \dots, \hat{y}_n)$ are the real and predicted values vectors).

$$\mathcal{L}(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{1}{n} \sum_{i=0}^n (\log(y_i + 1) - \log(\hat{y}_i + 1))^2 \quad (3.1)$$

The introduction of the logarithm makes MSLE penalize underestimates more than overestimates, introducing an asymmetry in the error curve. This property is related to the slope of the logarithmic curve [63] and provides the desired conservatism for the current application.

3.1.2. Minkowski logarithmic error

When the individual instances of ten-minute DEMs are aggregated into a long-term DEM, as in Eq. (2.2), we can see this is performed through a weighted Lebesgue, or L^p , space norm, a class of Banach spaces [64], where p coincides with the Wöhler exponent, m . Therefore, the length of the vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ in the n -dimensional real vector space \mathbb{R}^n can be described for L^p spaces as given by Eq. (3.2) [65].

$$\|\mathbf{x}\|^p = \left(\frac{1}{n} \sum_{i=0}^n |x_i|^p \right)^{1/p} \quad (3.2)$$

Thus, for the problem at hand, we can bring any timescale into the 10-min frame of reference through Eq. (2.2), a direct application of Eq. (3.2) for DEM with $p = m = 5$.

This equation is the point of departure for our physics-guided loss function, as it is the underlying physical principle which drives the long-term accumulation of DEMs. One could therefore reasonably place a complete focus on the long-term ability of the model to predict long-term, as seen in Eq. (3.3).

$$\mathcal{L}(\mathbf{Y}, \hat{\mathbf{Y}}) = \left(\sum_{i=0}^n |\log(y_i)|^p \right)^{1/p} - \left(\sum_{i=0}^n |\log(\hat{y}_i)|^p \right)^{1/p} \quad (3.3)$$

However, this approach would be detrimental, as it would mean that during the learning phase instead of comparing the L^p scaled residuals (and thus focusing on the estimation error), it would compare the residual of the L^p scaled total (long-term) measurements and estimations, which has no bearing on a ten-minute level. Therefore, if our loss function is to reflect the long-term DEM in a flexible way (i.e. not timescale-dependent), it must also remain able to estimate accurately at a ten-minute level.

Thus, based on the L^p norm and the logarithm function, and attempting to maintain a ten-minute level prediction accuracy, we introduce the Minkowski logarithmic error (MLE). This loss function can be

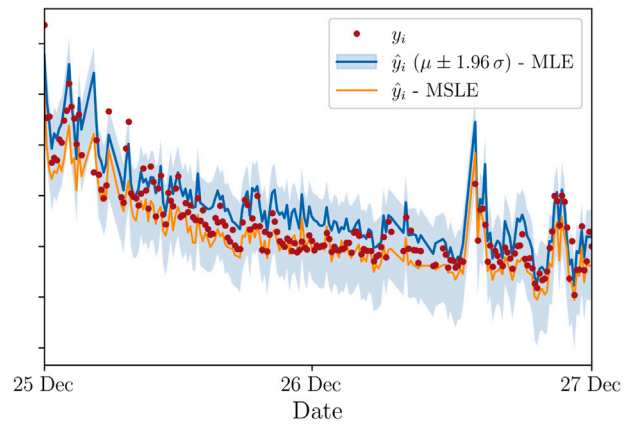


Fig. 4. Time series of measurements (y_i or DEM_{real} , in red) and predictions (\hat{y}_i or DEM_{pred} , in blue) with 95% confidence interval bounds (1.96σ , with σ standard deviation) of the relative error for OWT1 FA using MLE. Predictions using MSLE (\hat{y}_i in orange).

seen as an extension of the logarithmic loss function to any L^p metric, also known as Minkowski distance [66]. Eq. (3.4) describes this function mathematically, extended for $1 \leq p < +\infty$ in the n -dimensional vector space \mathbb{R}^n for $\mathbf{Y}, \hat{\mathbf{Y}}$. For our case the Wöhler exponent ($m = 5$) coincides with p .

$$\mathcal{L}(\mathbf{Y}, \hat{\mathbf{Y}}) = \left(\sum_{i=0}^n |\log(y_i + 1) - \log(\hat{y}_i + 1)|^p \right)^{1/p} \quad (3.4)$$

This loss function, as will be seen in Section 4, represents a fair compromise between ten-minute DEM and long-term DEM estimation and was implemented in the Keras framework. [67] provides an extensive introductory guide into custom loss functions – including physics-guided learning – for this specific TensorFlow API. A deeper discussion on the Minkowski Logarithmic Error and its mathematical development can be found in [68].

4. Results and discussion

As described in Section 3, the final architecture obtained through hyperparameter tuning is used during the training of the model using 50% of OWT1’s data (of the available 9 months). The final PINN model is used to estimate TW-TP ten-minute DEMs, which are then accumulated into a long-term DEM for each of the three turbines (both FA and SS) using nine months worth of data.

4.1. 10-minute level performance

We can begin inspecting the model’s performance on a ten-minute level by observing a two-day timeseries of the FA predictions with a 95% confidence interval for OWT1 using the Minkowski logarithmic error (MLE) loss function (MLE uses $p = m = 5$). The predictions using the mean squared logarithmic function (MSLE; control function) have also been added.

In Fig. 4 we can observe how all the measurements fall well within the 95% confidence interval. More specifically, we can see how the predictions are, on average, greater than the measurements. This is a direct result of the inclusion of the logarithmic function in the MLE, as it promotes over-prediction, introducing the desired conservatism to the model. When compared with the predictions using the MSLE control loss function, the overprediction incentive in these latter is not really noticeable.

We can further confirm that the model using MLE is generally over-predicting by plotting the predictions against the measurements (see Fig. 5) by smoothing their bivariate distribution using a kernel density

Table 2

Comparison of models' performance for MSLE and MLE (coefficient of determination, R^2 , RMSE expressed as a percentage of the mean DEM value, NRMSE (%)) and long-term DEM error, δ_{LT} , with $\delta_{LT} = 100 \times (\text{DEM}_{\text{real}} - \text{DEM}_{\text{pred}}) / \text{DEM}_{\text{real}}$ for the three turbine in each direction (fore-aft, FA, and side-to-side, SS). OW1 (T), training turbine.

	OWT1 (T)				OWT2				OWT3			
	FA		SS		FA		SS		FA		SS	
	MSLE	MLE	MSLE	MLE	MSLE	MLE	MSLE	MLE	MSLE	MLE	MSLE	MLE
R^2	0.85	0.87	0.95	0.98	0.86	0.85	0.94	0.98	0.85	0.88	0.94	0.98
NRMSE (%)	22.8	21.7	11.2	6.34	21.9	22.2	11.4	7.58	22.1	20.1	11.5	6.08
δ_{LT} (%)	7.99	1.25	9.01	2.21	6.95	-0.24	8.30	2.12	7.50	-0.87	9.38	2.83

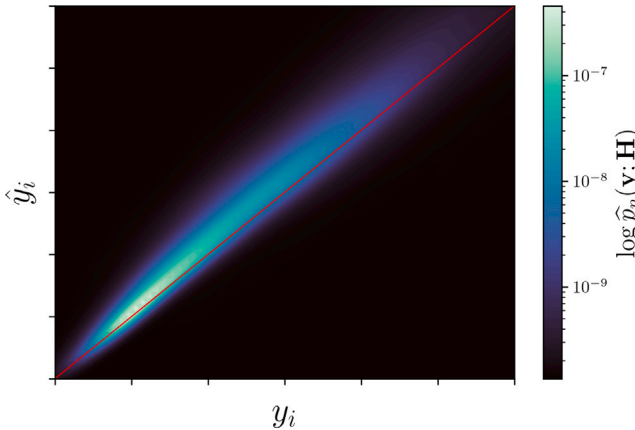


Fig. 5. Bivariate probability density plot of measurements (y_i or DEM_{real}) and predictions (\hat{y}_i or DEM_{pred}) with red line $y_i = \hat{y}_i$ for OWT1 FA using MLE. Colour scale is the log-normal distribution of the PDF $-\hat{p}_n(\mathbf{y}; \mathbf{H})$ with \mathbf{H} the 2×2 positive definite smoothing bandwidth matrix, selected through Scott's rule ($\sqrt{H_{ii}} = n^{-1/6} \sigma_i$) [71].

estimator (KDE) [69], wherein a multivariate gaussian kernel estimates the probability density function (PDF) [70] non-parametrically.

In Fig. 5, we can see how the majority of values are rather small when compared with the highest damage equivalent loads, being concentrated in the lower left corner, close to the origin (in light green/white). More interestingly, we can also see how there is a slight upward skew in the PDF distribution in relation to the identity line for $y_i = \hat{y}_i$ (measurements equal predictions). This shows us how there is a constant model overprediction ($\hat{y}_i > y_i$).

4.2. Long-term level performance

As mentioned in preceding sections, the major goal of this contribution lies with the accurate characterization of long-term DEMs through physics-guided learning by employing a custom-built loss function. In order to assess this, we can take a look at the long-term (nine months worth of data) results for the three turbines. This can be seen as a proof-of-concept of the fleet-leader model through cross-validation, opening the road for a farm-wide implementation. For this, we can inspect Table 2, the errors of TW-TP DEM estimation by direction for 10-min level through the coefficient of determination, R^2 and the root mean squared error, normalized as a percentage of the mean DEM value, where $NRMSE = 100 \cdot (1/n \sum_{i=1}^n (y_i - \hat{y}_i)^2)^{1/2} \cdot \bar{Y}$, but also for the long-term level (δ_{LT}) for the Minkowski logarithmic error (MLE) compare those control loss function, the mean squared logarithmic error (MSLE).

In Table 2 we can see how the introduction of a Φ -ML approach for the loss function greatly improves the models' performance on the long-term accumulations of Eq. (2.2) (δ_{LT}), whilst retaining and even slightly improving the ten-minute estimation performance (R^2 , NRMSE). All errors on long-term DEM estimation are well below 3%. These can be considered as excellent within the overall context of the substructure's design. There is usually an allowable error of similar magnitude between the substructure designer's and the wind turbine

designer's code. From real-world examples known to the authors, the results present in Table 2 are well below this industry allowable error. The ten-minute level improvement is more noticeable in SS, especially if we consider the NRMSE (4%–5% improvement). Indeed, the MLE loss function has not only meant improvements at 10-min and long-term levels (with the latter improving from up to 9.4% error to less than 3%), but has also prioritized overpredictions (a feature desired in order to ensure some conservatism), represented by a negative δ_{LT} , for OWT2 and 3 FA.

4.2.1. Side-to-side long-term error

However, despite the overall improvement of long-term performance due to MLE, the expressed objective of forcing overprediction (negative δ_{LT}), has not been attained for all SS cases and for OWT1 FA, with the long-term error being positive despite using the logarithmic in loss function to give incentive to overpredictions. In order to understand this, we inspect Fig. 6a, where we plot the predictions (\hat{y}_i) against the measurements (y_i) for DEM_{tl} of OWT 2. Additionally, we identify in purple the region (without balancing counterpart) where the model is severely underpredicting ($y_i \gg \hat{y}_i$).

Here, we can see that the model is overwhelmingly accurate, as the vast majority of values stick closely to the identity line (also given by the R^2 value in Table 2 of 0.98). However, there is indeed a region (in purple) where the measured values are much greater than the predicted values. We can additionally see how most of these values of the purple shaded region can be mapped by time-instances where there is a sudden variation in the yaw angle ($\partial\theta/\partial t_{i-1} > 10^\circ$, in orange).

In order to understand the rationale behind this mapping of the yaw transient, we can start by taking a look at Fig. 6b, where the lateral bending moment (M_{tl}) is plotted for an individual ten-minute time-instance located in the underpredicting region. Here, we can indeed see that there are two regions (identified in red), where there is a sudden variation (in this case, a drop) in the lateral bending moment ($\nabla M_{tl} \gg 0$) caused by a sudden variation in yaw. Therefore, we can understand that the model is unable to accurately capture these severe variations and, thus, underpredicts the fatigue loads. We can furthermore see how these sudden variations on M_{tl} are preceded by an equally sudden variation on the yaw angle (θ , in light blue in Fig. 6b). Thus, there is a link between yaw and lateral bending moment variation ($\partial\theta/\partial t_{i-1} \gg 0 \Rightarrow \nabla M_{tl} \gg 0$).

However the origin of these apparent large fatigue loads does not come from an actual sudden, physical variation in the SS loads. Instead it is due to one of the key assumptions of the global methodology [45] being violated. In general one assumes a unique value for the yaw angle for the whole of the ten-minutes. This assumption overwhelmingly holds true. But Fig. 6b shows that once this assumption breaks down, the bending moment calculation is not truthful to the physical reality (the bending moment variation regions in red in Fig. 6b do not have any physical meaning, they just are a result of the yaw angle not remaining approximately constant). Therefore, we are in the presence of a shortcoming of the decision to work with M_{tn} and M_{tl} bending moments and not of the machine learning model.

It could be argued that one should not work with M_{tl} and M_{tn} loads and use the strain gauges directly or use 1s SCADA for the yaw transformation. In Section 2.1 it was explained that working with

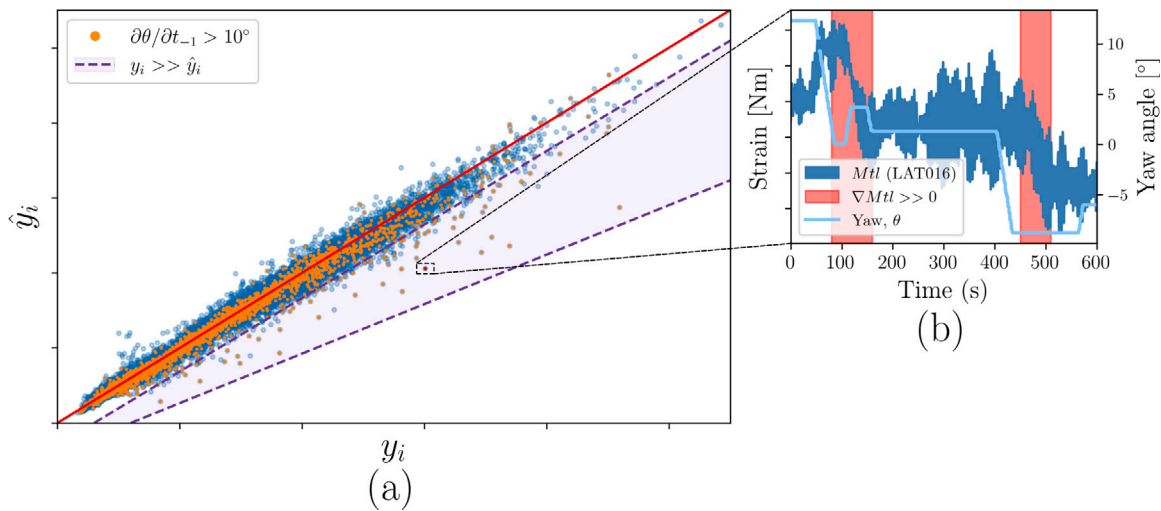


Fig. 6. (a) DEM predictions (\hat{y}_i) vs measurements (y_i). Red line identity ($y_i = \hat{y}_i$). Region of severe underprediction ($y_i \gg \hat{y}_i$) in purple. Time-instances of severe yaw angle transient ($\partial\theta/\partial t_{-1} > 10^\circ$) in orange. (b) Ten-minute timeseries window of lateral bending moment at LAT016 (M_{tl}) strain values (in Nm). Zones of sudden and severe variation of the lateral bending moment ($\nabla M_{tl} \gg 0$) identified in red. Yaw angle (θ) plotted in light blue.

M_{tl} and M_{in} made sense, as these are the ‘primary’ directions of the structure, with dynamics clearly being different between these two primary directions. When, alternatively, working with strain gauges directly, the two primary directions would be mixed and additional challenges to train a model would be posed. As for the use of 1s SCADA for the yaw angle, this would be challenging, as time-synchronization issues between the different systems would arise. Moreover, 1s SCADA is not readily available for all turbines/sites, which would hinder the widespread adoption of the current methodology.

4.2.2. Fore-aft long-term error

Similarly to Section 4.2.1, and as seen in Table 2, the issue of underprediction is not constrained merely to SS, but also for FA, with FA long-term error being positive (representing a global underprediction) for OWT1.

In order to understand the causes behind this behaviour, we resume the strategy of Section 4.2.1, this time by plotting $\|y_i - \hat{y}_i\|_5$, the L^5 norm of the errors, against the measurements for OWT1 FA, as shown in Fig. 7. Above this figure, we additionally plot the probability density function (PDF; see Fig. 5) and the cumulative distribution function (CDF), the integration of the PDF, $P_n(y) = \int_{-\infty}^x \hat{p}_n(y) dy$ [72].

In Fig. 7 we see the linear correlation between $\|y_i - \hat{y}_i\|_5$ (hereafter simply referred as errors) and y_i . If the errors are above the regression line (the residual is positive), then $y_i > \hat{y}_i$, which means that the model is underpredicting. As we can see, the mean residuals are above zero ($\bar{r}_i > 0$) to the right of the blue dashed line, representing less than 3% of the data. This can be further verified by observing the PDF and CDF plots, in which the measurements’ distribution is overwhelmingly concentrated on the lower band of the axis (to the left of the blue dashed line). This indicates that for the vast majority of fatigue loads the model is, as desired, overpredicting (left side of blue dashed line). However, the average of the residuals is dragged up due to the model’s behaviour for high load cases (which also represent a large role on the total fatigue life).

To further understand what is happening with the model in the high load case region we can plot the predictions (\hat{y}_i) against the measurements (y_i), as shown in Fig. 8a.

In Fig. 8a we can again see that the model’s performance is overwhelmingly accurate (most values are well within the 1.5σ confidence interval, CI, and closely follow the identity line, red). Therefore, the model is accurately able to capture the underlying behaviour of the turbine response, with it even overpredicting for small load values ($\hat{y}_i > y_i$ on bottom left corner of the figure). However, for a small subset

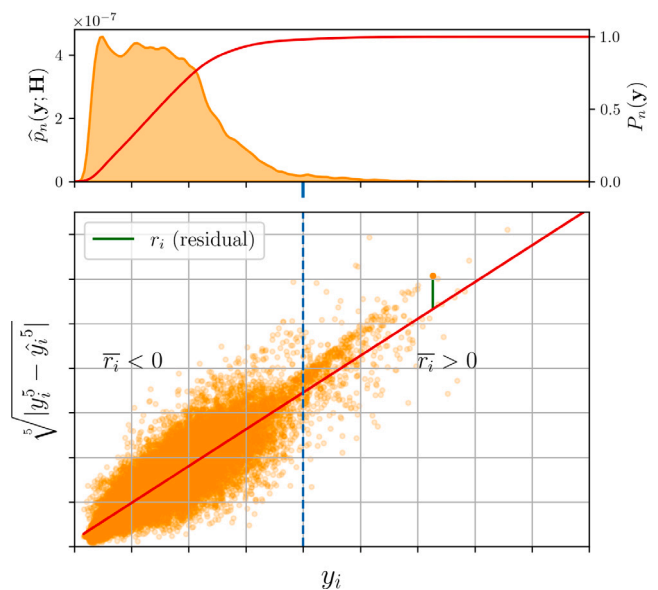


Fig. 7. L^5 norm of the errors, $\|y_i - \hat{y}_i\|_5$ vs measurements (y_i for DEM_n). r_i , residual, \bar{r}_i , mean of residuals and the red line represents the fitted linear regression of the distribution. Above, the PDF ($\hat{p}_n(y; \mathbf{H})$, orange) and CDF ($P_n(y)$, red) are plotted.

of data generally withstanding high fatigue loads, it is not capable of doing so (purple region). In the case of ten-minute predictions this is not as noticeable, as the R^2 for Table 2 demonstrates. However, when we accumulate and re-scale these errors, due to the presence of the Wöhler exponent ($m = 5$), these gain a higher preponderance on the overall long-term performance.

In the purple shaded region, the model is severely underpredicting ($y_i \gg \hat{y}_i$), and it does not have a overpredicting counterpart to cancel it out. If we take a look at an individual time-instance outlier (orange cross in Fig. 8a) we can better understand what is going on in this region. In Fig. 8b we can see for this time instance how the model’s DEM prediction (\hat{y}_i) is vastly inferior to the measurement (y_i) when plotted for the measured timeseries, when there is a drastic increase in DEM in a spike without the model being capable of accurately predicting this structural response. Further looking into Fig. 8c, where the rotational speed (ω , in RPM) of the rotor is plotted and we can

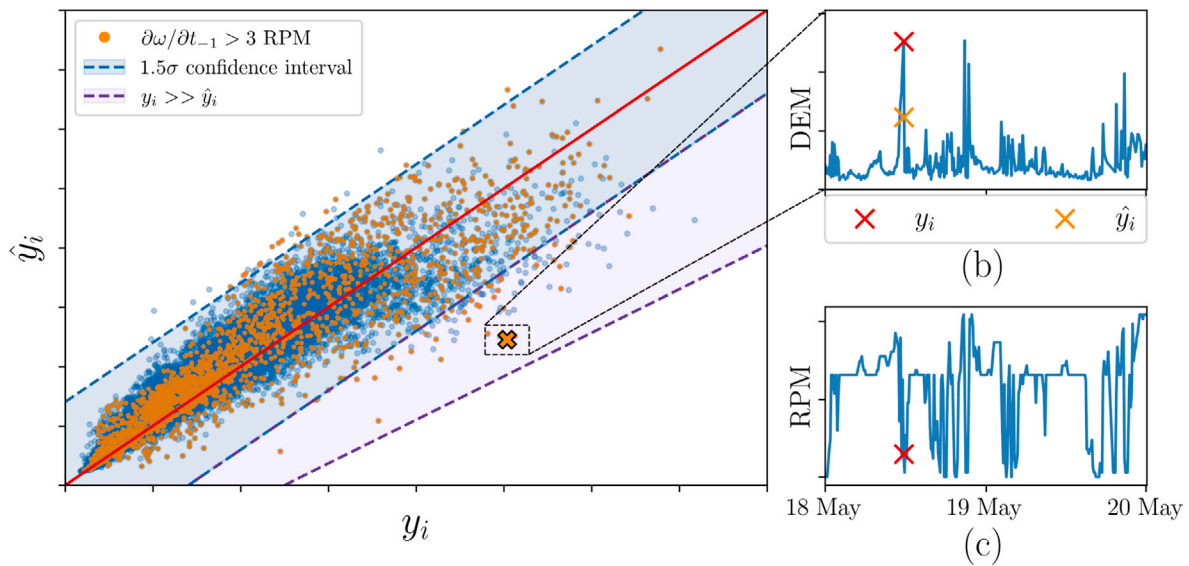


Fig. 8. (a) FA DEM (M_{in}) predictions (\hat{y}_i) vs measurements (y_i). Red line identity ($y_i = \hat{y}_i$). 1.5 σ confidence interval (86,6%) signalled in blue and zone of severe underprediction ($y_i \gg \hat{y}_i$) in purple. Time-instances of severe rotor rotational speed transient ($|\partial\omega/\partial t_{-1}| > 3$ RPM) in orange. (b) Outlier example (y_i , red cross) plotted for measured DEM time series (blue line) and predicted value for same time instance (\hat{y}_i , orange cross). (c) Same time series as (b), but for RPM.

identify the culprit for this underprediction: there is a severe and sudden variation on the RPM which causes a sudden spike in DEM.

If we define a threshold of 3 RPM for the transient (absolute variation considering the preceding time-instance, t_{-1}) of the rotational speed ($|\partial\omega/\partial t_{-1}| > 3$ RPM, a simple logical constraint), then we can return to Fig. 8a and map these time-instances by overlaying them (in orange). Indeed we can see that, when there is a severe and sudden variation of the rotor speed, albeit for some cases the overlap is not present, they are also able to correctly identify a good number of values within the purple region.

The sudden variation on RPM represents a tentative rotor stop/start. Thus, the PINN model appears not to be able to accurately replicate the turbine’s structural response when the turbine’s rotor suddenly stops. This is not wholly surprising as, due to the data quality and type, this phenomenon might not be captured by the data available.

4.3. Operational conditions

We have seen above how a Φ -ML approach greatly improves the model performance when DEMs are accumulated and re-scaled and the areas where it is lacking, due to the quality of the available data and some fundamental assumptions. In the present subsection, the ten-minute DEM estimation error (expressed as a percentage of the long-term DEM) is plotted for the mean squared logarithmic error (MSLE, control) and Minkowski logarithmic error (MLE) loss functions (for FA and SS) based on the operational cases: nominal (functioning turbine) or idling. We do this because the accurate portrayal of the fatigue loads faced during the different operating conditions a offshore wind turbine experiences has arguably become as important as a good general model. In modern monopile-foundation OWTs the complex interplay of aero-, hydro- and structural dynamics has meant that there is not one primary direction who is solely responsible for the fatigue incurred by the structure. This becomes specially relevant for idling conditions [33].

The performance of the MLE and MSLE models are presented in Fig. 9 for nominal and idling operational conditions for Fig. 9(a), fore-aft and Fig. 9(b), side-to-side.

In Fig. 9(a), we can see how the MLE loss function model has, for all turbines, an error distribution whose mean is negative for nominal operating conditions. This means that, on average, the model is overpredicting, as intended. As for the whiskers, we can detect an

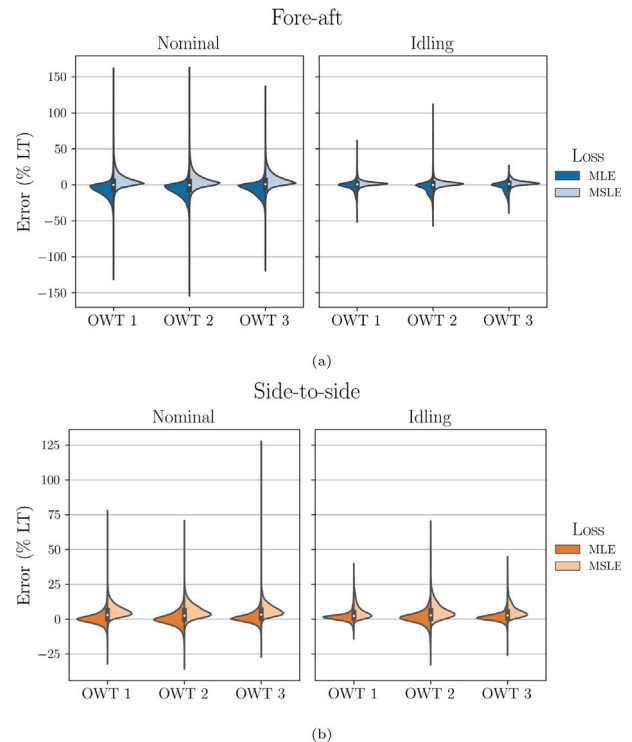


Fig. 9. Model prediction error based on loss function (Minkowski logarithmic error, MLE and mean squared logarithmic error, MSLE) expressed as a percentage of the long-term DEM for nominal and operational conditions for (a) fore-aft and (b) side-to-side. N.b. scales of (a) and (b) differ.

unbalance in OWT 1 (max above 150%, min below –150%), which can be reflected in the global long-term error being positive and thus indicating underprediction. As for the MSLE model, albeit the logarithm is also used, this is less evident. What is however noticeable is that the relative errors are much lower for idling, as fore-aft loads are lower under this operating condition. We can nevertheless still see how, specially for OWT 2 and 3, the MLE model is negatively skewed.

In Fig. 9(b) we can see how, for both nominal and idling conditions, the MLE model has its error distribution centred around zero, whereas

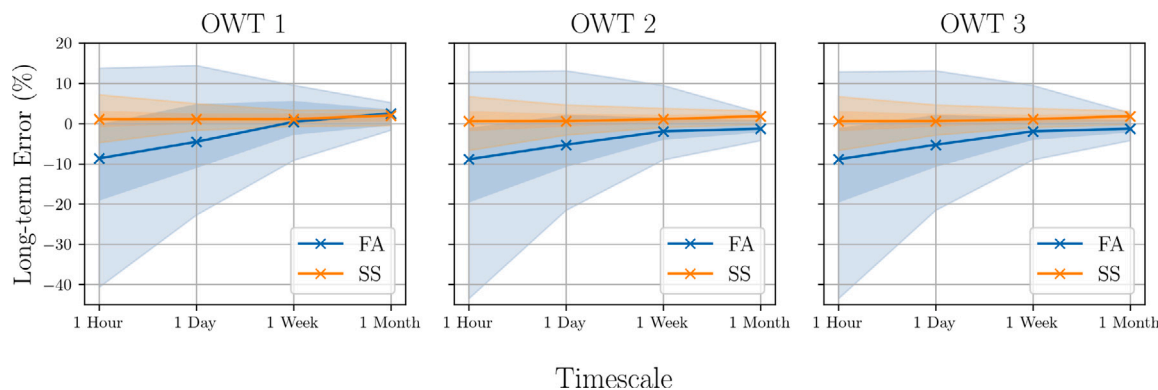


Fig. 10. Long-term DEM error (%) mean, 75% (darker blue) and 95% confidence interval bounds for different long-term DEM accumulation timescales for the three turbines.

the MSLE model has a positive bias. However, if we again take a look at the whiskers, these are greater on positive errors than negative ones.

Thus, globally, we can affirm that the MLE PINN is able to have a more accurate performance than a regular neural network approach with MSLE and better able to incorporate the logarithm-induced conservatism for all operating conditions.

4.4. Timescale

The main focus of this contribution lies with the accurate prediction of the long-term accumulation of DEMs. However, the long-term errors shown in Table 2 condense into a single value the performance of the PINN model, in which overpredictions and underpredictions on ten minute level will cancel each other out over 9 months of data. However, we can accumulated the ten-minute DEMs into different timescales (1 h, 1 day, 1 week, 1 month) and it is relevant to understand how the error propagates over these different timescales (see Fig. 10).

Fig. 10 shows that the spread on the error vastly reduces with the increasing timescale, rapidly converging, with it being centred around zero (this is to be expected, as negative and positive errors cancel each other out over time) for all turbines. We can observe, e.g., how for a monthly long-term DEM accumulation the error spread 95% confidence interval (1.96σ) is within a rather acceptable $\pm 5\%$. This has a practical relevance, as the error spread provides the confidence bounds operators must consider when comparing the accumulated DEMs over a given timescale. These results suggest that the method is suitable for monthly and even weekly fatigue accumulations, but might be insufficiently reliable for daily and hourly estimates. With SS estimates converging faster than those in FA direction. We can also observe how the incentive for overpredictions (reflected on the negative error mean) slowly converges with the increasing timescale towards zero, without actually reaching it (OWT 2 and 3) or slightly overpassing it (OWT 1). Interestingly enough, we can also observe for smaller timescales (1 hour/1 day), the long-term error seems to cap at around 15%. As for the SS error, we can see that its spread is much smaller and centred around zero (hovering slightly above it) for all timescales.

4.5. Fatigue damage accumulation

Finally, in Section 2.3 we have seen how the long-term DEM can be brought into a accumulated fatigue damage through the employment of Eq. (2.3). In this section, we show the results of the fore-aft and side-to-side fatigue damage accumulation for every month of the nine months of available data for the three turbines, whilst additionally presenting the progression of the relative error. We can see these on Fig. 11.

The figure shows the fatigue damage progressively increasing both fore-aft and side-to-side. We can see how, for OWT 2 and OWT 3 FA, the predicted damage slightly surpasses the measured damage (as intended

as per the inclusion of the safety-factor of the logarithm function in the Minkowski logarithmic error). The same cannot be said of OWT 1, due to the reasons explained in Section 4.2.2. Nevertheless, when we take a look at the relative error on the accumulated fatigue damage, it is kept within $\pm 5\%$ for all turbines.

More interestingly, we can see how, for all turbines, after the initial months where there is a strong variation, the relative error on the accumulated damage (δ_D , in green) begins to converge to a constant value (after roughly 6+ months). This is consistent with [45], where the same 9–12 months timeframe was identified until convergence. The convergence on the error is essential if the developed model is to be used in prognostic applications, extrapolating its predictions into the future.

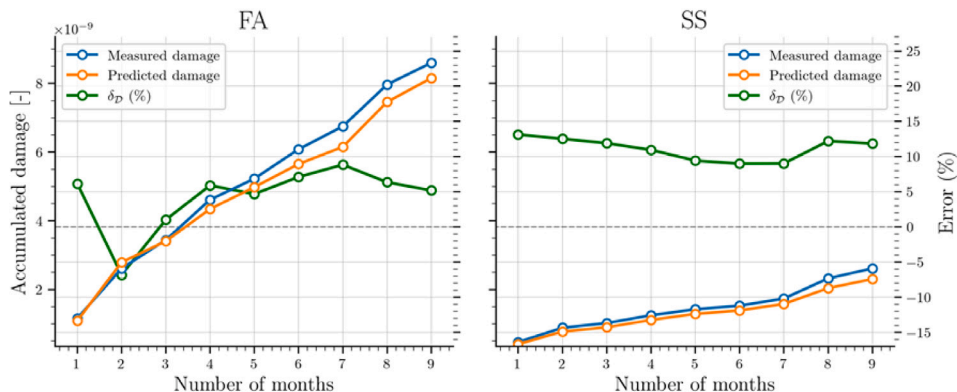
This variation on the error is less noticeable in the side-to-side direction, but there is nevertheless, a stability. As seen in Section 4.2.1, the relative error on the accumulated fatigue damage will be positive (between 10 and 15%), because the model is underpredicting (measured damage is higher than the predicted). This might seem inconsistent with Table 2, which showed SS long-term DEM errors that are below 3%. However, accounting for power of m in Eq. (2.3) means a 3% error in DEM_{LT} results in a 15% error in fatigue damage. The power of m is also the reason why the ratio between FA and SS increases: for DEMs this ratio (taken every month) is between 1.20 and 1.34, whereas for damage it is between 2.48 and 4.32 (1.20^5 and 1.34^5 , respectively).

5. Conclusions and future work

In this contribution, we propose and analyse the performance of a data-based methodology formulated through physics-guided learning for neural networks. This is used in a fleet-leader strategy for tower-monopile transition piece interface fatigue loads long-term estimation. The explicit objective of long-term estimation is stated by explaining the rationale and procedure behind the accumulation of ten-minute instances of damage equivalent moments (DEM) into a long-term DEM. This metric is further brought into a fatigue damage value. The knowledge required for the accumulation of 10-min DEM (by employing a Lebesgue norm) is additionally used to shape a custom loss function – the Minkowski logarithmic error (MLE) – which incorporates physical knowledge of the problem and the properties of the logarithmic function, forcing the model to slightly overpredict and therefore adding an element of conservativeness.

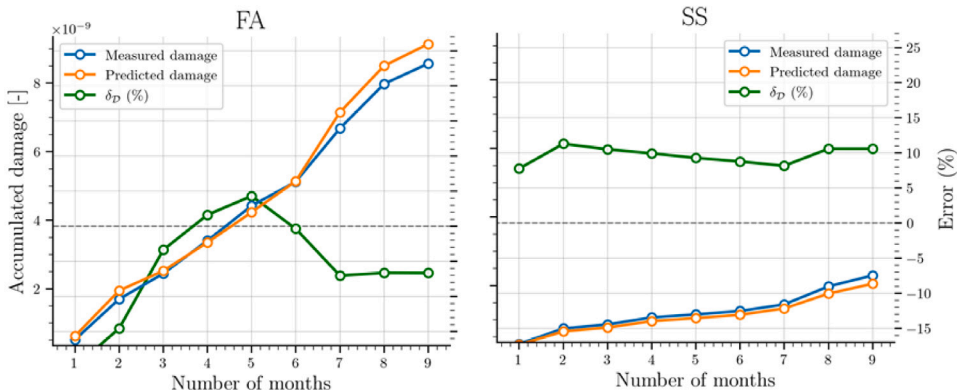
This approach involved the use of nine months of monitored data in three real-world locations consisting of ten-minute statistics of SCADA, wave, acceleration (inputs) and strain data (used as the target for the neural network model). The available input data statistics were reduced to a smaller number of explanatory features through the employment of recursive feature elimination and the optimal neural network model topology found through Bayesian hyperparameter optimization.

OWT 1



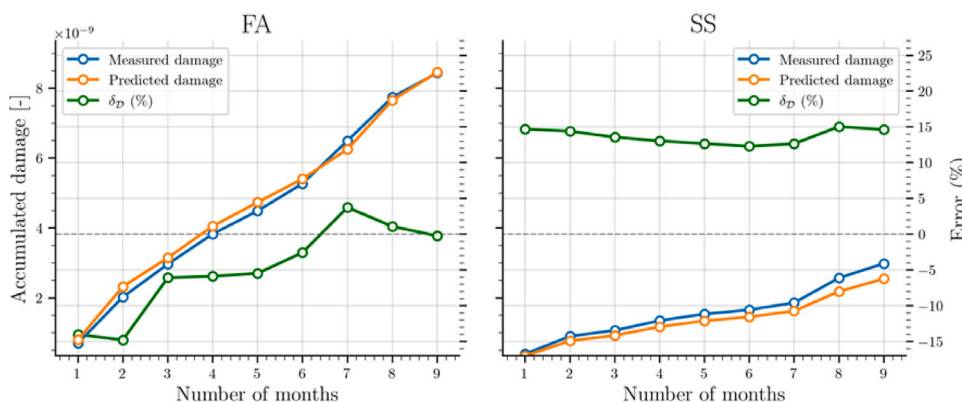
(a)

OWT 2



(b)

OWT 3



(c)

Fig. 11. Accumulated measured (blue) and predicted (orange) fore-aft fatigue damage based on number of months. Relative error on damage, δ_D , on right-hand axis with $\delta_D = 100 \times (D_{real} - D_{pred})/D_{real}$ for OWT 1, 2 and 3 (a), (b) and (c), respectively.

In the results section it was seen how the employment of the Minkowski logarithmic error is able to accurately predict fore-aft and side-to-side DEM timeseries on a ten-minute basis. Furthermore, the MLE was favourably compared with a control loss function (mean squared logarithmic error), with performance greatly increasing for long-term DEM estimation. It was additionally noted that, for some

cases, the model was still underpredicting (and therefore, contradicting the rationale behind the use of the logarithm). The reason behind this behaviour was scrutinized by inspecting the residuals and discovering that, for the side-to-side direction, due to some fundamental assumptions on the bending moment calculation, the model was presenting errors without an underlying physical explanation. The same analysis

was done for the fore-aft direction, where it was seen that rotor stops were not wholly captured by the model as it lacked data.

Additionally, the performance of the model was inspected in relation to nominal and idling operational conditions, with a greater performance over the board for the Minkowski logarithm error when compared with the control loss function. Thereafter, the influence of the timescale – hourly, daily, weekly and monthly accumulation – on the long-term DEM prediction accuracy was inspected, with the errors balancing each other out (centred around zero) and the spread rapidly converging to towards zero. Finally, the long-term DEM was translated into fatigue damage and its progressive accumulation and error convergence verified. It was seen how these values compare favourably in relation to industry practice.

The next logical step to be taken in this research is to apply this fleet-leader model in a full farm-wide setting, and not just for turbines instrumented with strain gauges where cross-validation is possible. Such an implementation can be seen as pilot in what is hoped to become a commonplace practice among wind farm operators in the future. This is naturally dependent on farm-wide data availability, but, in the case of the farm used for this study, farm-wide SCADA and acceleration data will become a reality in the very near future, with sensor installation already rolling out (at the time of writing). Coupled with the farm-wide implementation, and in order to quantify the uncertainty related to the predictive model and therefore, introduce a metric of confidence one might have on said model, future research will also focus on applying a neural network predicated upon a Bayesian probabilistic framework to a farm-wide setting. The main advantage of such a Bayesian neural network approach is the probabilistic output (in the form of the mean and standard deviation of predicted DELs), which can be leveraged to indicate the model's confidence in its output through the coefficient of variation.

Finally, a study quantifying event-related fatigue is also necessary. If wind farm operators are to effectively understand the fatigue life consumption of their assets and leverage this knowledge next to the grid operator, they are required to quantify fatigue life consumption of demand-induced events, such as stopping or curtailing rotor operation due to grid requirements. The accurate translation of these events into a fatigue consumption framework and further into an economical impact metric through a value of information analysis is of the utmost importance for a fair and informed negotiation between all wind energy stakeholders.

In summation, in this contribution, it was shown how problem-specific knowledge employed in physics-guided learning of neural networks greatly increases model performance for multi-timescale objectives, the data quality dependant pitfalls and their probable solutions and how the marriage between engineering knowledge and dedicated sensor data are essential for modern structural health assessments in offshore wind support structures.

Declaration of competing interest

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

Data availability

The authors do not have permission to share data.

Acknowledgements

The research carried out in this contribution was financed by the ICON Supersized 4.0 project. The authors would additionally like to acknowledge 24SEA and Parkwind for the data collection and management.

References

- [1] Iván Pineda, Pierre Tardieu (Eds.), The European offshore wind industry: Key trends and statistics 2016, 2022, <https://windeurope.org/intelligence-platform/product/the-european-offshore-wind-industry-key-trends-and-statistics-2016/>, (Accessed 15 August 2022).
- [2] Rory O'Sullivan (Ed.), Wind energy in Europe: 2021 statistics and the outlook for 2022–2026, 2022, <https://windeurope.org/intelligence-platform/product/wind-energy-in-europe-2021-statistics-and-the-outlook-for-2022-2026/>, (Accessed 15 August 2022).
- [3] Byron W Byrne, Ross A McAdam, Harvey J Burd, Guy T Houlby, Chris M Martin, William JA P Beuckelaers, L Zdravkovic, DMG Taborda, DM Potts, RJ Jardine, et al., PISA: new design methods for offshore wind turbine monopiles, in: Offshore Site Investigation Geotechnics 8th International Conference Proceeding, 142, (161) Society for Underwater Technology, 2017, pp. 142–161.
- [4] Subhamoy Bhattacharya, Challenges in design of foundations for offshore wind turbines, Eng. Technol. Ref. 1 (1) (2014) 922.
- [5] Francisco de N Santos, Koen Robbelein, Pietro D'Antuono, Nymfa Noppe, Wout Weijtjens, Christof Devriendt, Towards a fleetwide data-driven lifetime assessment methodology of offshore wind support structures based on SCADA and SHM data, in: Piervincenzo Rizzo, Alberto Milazzo (Eds.), European Workshop on Structural Health Monitoring, in: Lecture Notes in Civil Engineering, vol. 253, Springer International Publishing, Cham, ISBN: 978-3-031-07254-3, 2023, pp. 123–132.
- [6] R. Miedema, Offshore Wind Energy Operations & Maintenance Analysis, Hoge School Van Amsterdam, 2012.
- [7] Maria Martinez-Luengo, Mahmood Shafiee, Guidelines and cost-benefit analysis of the structural health monitoring implementation in offshore wind turbine support structures, Energies 12 (6) (2019) 1176.
- [8] Xiaoxu Liu, Zhiwei Gao, Michael Z.Q. Chen, Takagi-Sugeno fuzzy model based fault estimation and signal compensation with application to wind turbines, IEEE Trans. Ind. Electron. 64 (7) (2017) 5678–5689.
- [9] Zhiwei Gao, Xiaoxu Liu, An overview on fault diagnosis, prognosis and resilient control for wind turbine systems, Processes 9 (2) (2021) 300.
- [10] Jannie Jessen Nielsen, John Dalsgaard Sørensen, On risk-based operation and maintenance of offshore wind turbine components, Reliab. Eng. Syst. Saf. 96 (1) (2011) 218–229.
- [11] Marcel Schedat, Torsten Faber, Abiman Sivanesan, Structural health monitoring concept to predict the remaining lifetime of the wind turbine structure, in: 2016 International Conference on the Domestic Use of Energy, DUE, IEEE, 2016, pp. 1–5.
- [12] Alexandros Iliopoulos, Wout Weijtjens, Danny Van Hemelrijck, Christof Devriendt, Fatigue assessment of offshore wind turbines on monopile foundations using multi-band modal expansion, Wind Energy 20 (8) (2017) 1463–1479.
- [13] Lisa Ziegler, Ursula Smolka, Nicolai Cosack, Michael Muskulus, Brief communication: Structural monitoring for lifetime extension of offshore wind monopiles: can strain measurements at one level tell us everything? Wind Energy Sci. 2 (2) (2017) 469–476.
- [14] Nymfa Noppe, Alexandros Iliopoulos, Wout Weijtjens, Christof Devriendt, Full load estimation of an offshore wind turbine based on SCADA and accelerometer data, 753, (7) IOP Publishing, 2016, 072025.
- [15] L. Vera-Tudela, M. Kühn, Evaluation of a wind turbine fatigue load monitoring system based on standard SCADA signals in different wind farm flow conditions, in: Dewek 2015, 49, (441) 2015.
- [16] Artur Movsessian, Marcel Schedat, Torsten Faber, Modelling tower fatigue loads of a wind turbine using data mining techniques on SCADA data, Wind Energy Sci. Discuss. (2020) 1–20.
- [17] Nymfa Noppe, Clemens Hübler, Christof Devriendt, Wout Weijtjens, Validated extrapolation of measured damage within an offshore wind farm using instrumented fleet leaders, 1618, IOP Publishing, 2020, 022005.
- [18] Francisco de N Santos, Nymfa Noppe, Wout Weijtjens, Christof Devriendt, Data-driven farm-wide fatigue estimation on jacket-foundation OWTs for multiple SHM setups, Wind Energy Sci. 7 (1) (2022) 299–321.
- [19] Michel Géradin, Daniel J. Rixen, Mechanical Vibrations: Theory and Application To Structural Dynamics, John Wiley & Sons, 2014.
- [20] Ursula Smolka, Po Wen Cheng, et al., On the design of measurement campaigns for fatigue life monitoring of offshore wind turbines, in: The Twenty-Third International Offshore and Polar Engineering Conference, International Society of Offshore and Polar Engineers, 2013.
- [21] Luis Vera-Tudela, Martin Kühn, On the selection of input variables for a wind turbine load monitoring system, Proc. Technol. 15 (2014) 726–736.
- [22] Luis Vera-Tudela, Martin Kühn, Analysing wind turbine fatigue load prediction: The impact of wind farm flow conditions, Renew. Energy 107 (2017) 352–360.
- [23] Laura Schröder, Towards digital twin technology: Wind farm operation analysis and optimization using model-supported data analytics (Ph.D. thesis), Danmarks Tekniske Universitet, 2020.
- [24] Luis David Avendaño-Valencia, Imad Abdallah, Eleni Chatzi, Virtual fatigue diagnostics of wake-affected wind turbine via Gaussian process regression, Renew. Energy 170 (2021) 539–561.

- [25] Charilaos Mylonas, Imad Abdallah, E.N. Chatzi, Deep unsupervised learning for condition monitoring and prediction of high dimensional data with application on windfarm scada data, in: *Model Validation and Uncertainty Quantification*, Volume 3, Springer, 2020, pp. 189–196.
- [26] Charilaos Mylonas, Imad Abdallah, Eleni Chatzi, Conditional variational autoencoders for probabilistic wind turbine blade fatigue estimation using supervisory, control, and data acquisition data, *Wind Energy* 24 (10) (2021) 1122–1139.
- [27] Charilaos Mylonas, Eleni Chatzi, Remaining useful life estimation for engineered systems operating under uncertainty with causal GraphNets, *Sensors* 21 (19) (2021) 6325.
- [28] Artur Movsessian, Marcel Schedat, Torsten Faber, Feature selection techniques for modelling tower fatigue loads of a wind turbine with neural networks, *Wind Energy Sci.* 6 (2) (2021) 539–554.
- [29] Luis David Avendano-Valencia, Eleni N Chatzi, Dmitri Tcherniak, Gaussian process models for mitigation of operational variability in the structural health monitoring of wind turbines, *Mech. Syst. Signal Process.* 142 (2020) 106686.
- [30] Baran Yeter, Yordan Garbatov, Structural integrity assessment of fixed support structures for offshore wind turbines: A review, *Ocean Eng.* 244 (2022) 110271.
- [31] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, Liu Yang, Physics-informed machine learning, *Nat. Rev. Phys.* 3 (6) (2021) 422–440.
- [32] J Gengenbach, K Mikkelsen, F Rudinger, M Brommundt, JS Grelund, Design challenges of XL monopiles, in: *Proceedings of the EWEA (European Wind Energy Association) Offshore*, 2015.
- [33] Francisco de N Santos, Nymfa Noppe, Wout Weijtjens, Christof Devriendt, Results of fatigue measurement campaign on XL monopiles and early predictive models, *J. Phys. Conf. Ser.* 2265 (2022) 032092.
- [34] Elizabeth J Cross, SJ Gibson, MR Jones, DJ Pitchforth, S Zhang, TJ Rogers, Physics-informed machine learning for structural health monitoring, in: *Structural Health Monitoring Based on Data Science Techniques*, Springer, 2022, pp. 347–367.
- [35] Maziar Raissi, Paris Perdikaris, George E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* 378 (2019) 686–707.
- [36] Vicente Negro, José-Santos López-Gutiérrez, M Dolores Esteban, Pablo Alberdi, Mario Imaz, José-María Serracalera, Monopiles in offshore wind: Preliminary estimate of main dimensions, *Ocean Eng.* 133 (2017) 253–261.
- [37] Victor Igwemezie, Ali Mehmanparast, Athanasios Kolios, Current trend in offshore wind energy sector and material requirements for fatigue resistance improvement in large wind turbine support structures—A review, *Renew. Sustain. Energy Rev.* 101 (2019) 181–196.
- [38] Joey Velarde, Claus Kramhöft, John Dalsgaard Sørensen, Gianluca Zorzi, Fatigue reliability of large monopiles for offshore wind turbines, *Int. J. Fatigue* 134 (2020) 105487.
- [39] Ravi Pandit, Davide Astolfi, Jiarong Hong, David Infield, Matilde Santos, SCADA data for wind turbine data-driven condition/performance monitoring: A review on state-of-art, challenges and future trends, *Wind Eng.* (2022) 0309524X221124031.
- [40] Meetnet vlaamse banken, 2022, <https://meetnetvlaamsebanken.be/>, (Accessed 20 May 2022).
- [41] Turan Dirlik, Application of computers in fatigue analysis (Ph.D. thesis), University of Warwick, 1985.
- [42] Gabriel Marsh, Colin Wignall, Philipp R Thies, Nigel Barltrop, Atila Incecik, Vengatesan Venugopal, Lars Johanning, Review and application of rainfall residue processing techniques for accurate fatigue damage estimation, *Int. J. Fatigue* 82 (2016) 757–765.
- [43] J.J. Kaulzarich, The palmgren-miner rule derived, in: *Tribology Series*, vol. 14, Elsevier, 1989, pp. 175–179.
- [44] Lisa Ziegler, Michael Muskulus, Comparing a fracture mechanics model to the SN-curve approach for jacket-supported offshore wind turbines: Challenges and opportunities for lifetime prediction, in: *International Conference on Offshore Mechanics and Arctic Engineering*, vol. 49972, American Society of Mechanical Engineers, 2016, V006T09A054.
- [45] Clemens Hübler, Wout Weijtjens, Raimund Rolfes, Christof Devriendt, Reliability analysis of fatigue damage extrapolations of wind turbines using offshore strain measurements, *J. Phys. Conf. Ser.* 1037 (3) (2018) 032035.
- [46] H.B. Hendriks, B.H. Bulder, Fatigue Equivalent Load Cycle Method, Netherlands: Energy Research Foundation, 1995.
- [47] Nicolai Cosack, Fatigue load monitoring with standard wind turbine signals (Ph.D. thesis), Universität Stuttgart, 2010.
- [48] Marc Seidel, Sven Voormeeren, Jan-Bart van der Steen, State-of-the-art design processes for offshore wind turbine support structures: Practical approaches and pitfalls during different stages in the design process, *Stahlbau* 85 (9) (2016) 583–590.
- [49] DNVGL, Fatigue design of offshore steel structures, Recommended practice RP-C203, 2016.
- [50] Francisco de N Santos, Nymfa Noppe, Wout Weijtjens, Christof Devriendt, Two-tier model for wind turbine fatigue assessment based on SCADA-dependent neural networks, in: *10th International Conference on Structural Health Monitoring of Intelligent Infrastructure*, 2021.
- [51] E. Keany, BorutaShap: A wrapper feature selection method which combines the boruta feature selection algorithm with Shapley values, 2020, Zenodo <https://zenodo.org/record/4247618> (Accessed October 25 2021).
- [52] Isabelle Guyon, Jason Weston, Stephen Barnhill, Vladimir Vapnik, Gene selection for cancer classification using support vector machines, *Mach. Learn.* 46 (1) (2002) 389–422.
- [53] Francisco de N Santos, Nymfa Noppe, Wout Weijtjens, Christof Devriendt, Input parameter selection for full load damage neural network model on offshore wind structures, in: *Proceedings of 16th EAWC PhD Seminar on Wind Energy*, 2020.
- [54] James Bergstra, Rémi Bardenet, Yoshua Bengio, Balázs Kégl, Algorithms for hyper-parameter optimization, *Adv. Neural Inf. Process. Syst.* 24 (2011).
- [55] Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, Ameet Talwalkar, Hyperband: A novel bandit-based approach to hyperparameter optimization, *J. Mach. Learn. Res.* 18 (1) (2017) 6765–6816.
- [56] Jonas Mockus, Bayesian Approach to Global Optimization: Theory and Applications, vol. 37, Springer Science & Business Media, 2012.
- [57] Christopher K. Williams, Carl Edward Rasmussen, Gaussian Processes for Machine Learning, vol. 2, MIT press Cambridge, MA, 2006.
- [58] Tom O'Malley, Elie Bursztein, James Long, François Chollet, Haifeng Jin, Luca Invernizzi, et al., KerasTuner, 2019, <https://github.com/keras-team/keras-tuner>.
- [59] Xavier Glorot, Antoine Bordes, Yoshua Bengio, Deep sparse rectifier neural networks, in: *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings*, 2011, pp. 315–323.
- [60] Dan Hendrycks, Kevin Gimpel, Gaussian error linear units (gelus), 2016, arXiv preprint arXiv:1606.08415.
- [61] Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al., The elements of statistical learning, vol. 1(10), Springer series in statistics New York, 2001.
- [62] Jared Willard, Xiaowei Jia, Shaoming Xu, Michael Steinbach, Vipin Kumar, Integrating scientific knowledge with machine learning for engineering and environmental systems, *ACM Comput. Surv.* (2021).
- [63] Timothy Gowers, The exponential and logarithmic functions, in: Timothy Gowers, June Barrow-Gree, Imre Leader (Eds.), *The Princeton Companion to Mathematics*, Princeton University Press, 2008, pp. 199–202.
- [64] Nicolas Bourbaki, Sur certains espaces vectoriels topologiques, in: *Annales de L'Institut Fourier*, vol. 2, 1950, pp. 5–16.
- [65] Michel Marie Deza, Elena Deza, Encyclopedia of distances, in: *Encyclopedia of Distances*, Springer, 2009, pp. 1–583.
- [66] John M. Lee, Riemannian Manifolds: An Introduction to Curvature, vol. 176, Springer Science & Business Media, 2006.
- [67] Imme Ebert-Uphoff, Ryan Lagerquist, Kyle Hilburn, Yoonjin Lee, Katherine Haynes, Jason Stock, Christina Kumler, Jebb Q Stewart, CIRA guide to custom loss functions for neural networks in environmental sciences—Version 1, 2021, arXiv preprint arXiv:2106.09757.
- [68] Francisco de N Santos, Pietro D'Antuono, Nymfa Noppe, Wout Weijtjens, Christof Devriendt, Minkowski logarithmic error: A physics-informed neural network approach for wind turbine lifetime assessment, in: *30th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2022, European Symposium on Artificial Neural Networks*, 2022, pp. 357–362.
- [69] Yen-Chi Chen, A tutorial on kernel density estimation and recent advances, *Biostat. Epidemiol.* 1 (1) (2017) 161–187.
- [70] Emanuel Parzen, On estimation of a probability density function and mode, *Ann. Math. Stat.* 33 (3) (1962) 1065–1076.
- [71] David W. Scott, Multivariate Density Estimation: Theory, Practice, and Visualization, John Wiley & Sons, 2015.
- [72] Kun Il Park, M. Park, Fundamentals of Probability and Stochastic Processes with Applications to Communications, Springer, 2018.