Fleetwide Interface Fatigue Load Prediction of an Operational Offshore Wind Farm Using a Single Accelerometer and SCADA Data

FRANCISCO DE N. SANTOS, NYMFA NOPPE,
WOUT WEIJTJENS and CHRISTOF DEVRIENDT

ABSTRACT

The operational life of offshore wind turbines is in part driven by the fatigue life of key structural components, such as the substructure. In recent years, fatigue life management of operational assets has become evermore important, as older farms are closing on their design lifetime and newer farms are designed with tighter margins. To support decisions on fatigue lifetime, it is advantageous to monitor the fatigue progression in these structures through SHM. However, a full instrumentation of every asset in a farm to assess the fatigue life of the substructure is considered economically infeasible.
The drive for SHM in wind has been accompanied by the increased availability of intelligent data-driven methodologies which have attempted to provide the same information without the need for additional hardware. One such cost-effective approach, as described in [1], uses the available supervisory control and data acquisition (SCADA) systems, coupled with acceleration measurements to predict the fatigue life of an offshore wind turbine. The use of acceleration measurement data has been proven critical for capturing the complex dynamics of offshore wind turbines. In this contribution, we present the results of said data-driven approach using SCADA and Internet of Things (IoT) accelerometer installed at nacelle-level to monitor the fatigue life for the entirety of a real-world offshore wind farm comprised of 23 turbines, with a specific focus on long-term damage equivalent fatigue loads (DEL) estimation [2]. The availability of acceleration measurements for all locations in particular, is fundamental, as to cover all possible differences in natural frequencies between turbines would be nearly impossible if solely relying on wave and tidal data. To achieve this goal, a neural network architecture is used, enhanced by physics-informed learning focused on long-term estimation, trained and validated on so-called fleet-leader (three turbines instrumented with strain gauges, which provide the ground truth). In this study, we pay special attention to the farm-wide validation, cross-validation and extrapolation of these models as well as the performance for different operational conditions. Finally, this study is undertook for a real-world instrumentation setup, unparalleled in its scale, and can thus be indicative of future trends for SHM in offshore wind: farm-wide instrumentation and monitoring of structural health based on acceleration measurements, enabling a greater trustworthiness on reliability and durability estimation over the serviceable lifetime.

Francisco de N Santos, PhD Student, Email: francisco.de.nolasco.santos@vub.be. OWI-Lab, Acoustics and Vibrations Research Group (AVRG), Toegepasta Mechanica, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

INTRODUCTION

With concerns regarding proper and sustainable asset management increasing, along with older wind farms honing in on their design lifetime and newer farms being designed with smaller margins, the fatigue life management of offshore wind turbines has become evermore important in the current industrial and research landscape. Specifically, the monitoring of the fatigue progression through structural health monitoring (SHM), namely on the substructure, is increasingly commonplace among operators as an additional supporting tool for informed fatigue lifetime decisions. Keeping tabs on the fatigue life of offshore wind turbines is particularly important, as these structures are designed for dynamics rather than for bearing capacity [3]. Therefore, in order to replace visual inspections, which can be dangerous, time consuming and costly, remote, non-intrusive health monitoring systems have been developed [4]. With fatigue being a long time-scale type of damage [5], the sensor selection has traditionally lied with strain gauge instrumentation [6, 7]. However, farm-wide strain gauge installation lacks economical viability, and alternatives using supervisory control and data acquisition (SCADA) coupled with acceleration data for farm-wide fatigue load estimation have appeared [8]. The use of acceleration measurement data in particular, has been proven critical to capture the complex dynamics of offshore wind turbines [9].
Nevertheless, despite diverse data-oriented (based on long-term SCADA and accelerations) methodologies for fatigue load estimation on offshore wind turbines' substructures (from artificial neural networks [10], to Gaussian processes regression [11], graph neural networks [12], and physics-guided neural networks [1]), all hereto presented approaches have still been in the proof-of-concept stage as, at most, only a reduced number of real-world turbines of a farm have been sufficiently instrumented. In this contribution we present - to the best of the authors' knowledge - the first truly farm-wide implementation of a long-term fatigue load estimation (in the form of damage equivalent loads, DEL) data-driven model. We specifically pay attention to the validation, cross-validation and extrapolation of the trained models, both for fore-aft, as for side-to-side long-term DEL estimation (based on three months of real-world training data).

METHODOLOGY AND DATA

In the current contribution, three months of data (December 2022 – February 2023) from a real-world offshore wind farm with XL monopile foundations (9.5 MW turbines and water depths of up to 36 m) located in the Belgian North Sea was used. More specifically, SCADA data and accelerations from a nacelle-installed *'IoT'* accelerometer were collected for all locations within the farm and, for three locations (WT11, WT14, WT18) – representative of the different seabed-depth clusters present in the farm –, strain gauges installed along the inner circumference at the tower-transition piece interface level. The full dataset description can be found in Table I.

The presence of strain gauges on three of the turbines is relevant because, albeit the farm-wide strain gauge instrumentation is cost-prohibitive, as discussed in the introduction, whatever data-driven model that uses as input SCADA and accelerations, will still require the knowledge of a ground truth (a damage-sensitive feature, only attainable through strain gauge instrumentation) for the training locations before farm-wide extrap-

TABLE I: DESCRIPTION OF DATASETS FROM THE MEASUREMENT CAMPAIGN. NOTE THE CONFLICTING SAMPLING FREQUENCIES. EACH DATA-TYPE IS PROCESSED INTO 10-MINUTE 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.			
	Accelerometer	12.5 Hz	FA acceleration	g
			SS acceleration	g
			Z acceleration	g
	Target statistics (10-min): mean, minimum, maximum, root mean square.			
Target	Strain gauges	30 Hz	Normal bending moment (M_{tn})	MNm
			Lateral bending moment (M_{tl})	MNm
	Target statistics (10-min): damage equivalent loads ($\mathbf{DEL_{tn}}$ and $\mathbf{DEL_{tl}}$).			

olation. The strain gauges are therefore relevant, as they enable the calculation of the normal and lateral bending moments (M_{tn} and M_{tl}, respectively). Then, by employing a rainflow counting algorithm [13, 14], holding the linear damage accumulation hypothesis as true (Palmgren-Miner's rule) [15], and through the employment of the Wöhler exponent (the negative inverse slope of the SN curve) [16], a damage-sensitive feature, the damage equivalent loads (DEL), can be calculated for any ten minute window [17]. Therefore, given the Wöhler exponent, m, the number of cycles, n_i, of a given stress range, σ_i and the tower-transition piece outer and inner radii, r_o and r_i, respectively at strain gauge locations we can calculate the DEL, as given by Equation 1 [18]. 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.

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

The DELs are calculated for both the FA (DEL_{tn}) and SS (DEL_{tl}) directions. However, as discussed in [1], accuracy on DEL estimation at a ten-minute level isn't a sufficient condition to determine a model's success: one must also be able to accumulate DELs on a longer timeframe. Thus, by Palmgren-Miner's rule, we can further combine n equivalent load ranges that have been derived for the same reference cycle number and Wöhler exponent through the m-root of the weighted summation of the m-power DEL instance [19], as seen in Equation 2. Here DEL_{LT} represents the long-term

DEL accumulation of i ten-minute DEL instances. In this equation every i ten-minute time-instance DEL represents a damage load with identical occurrence probability of $1/n$ [20].

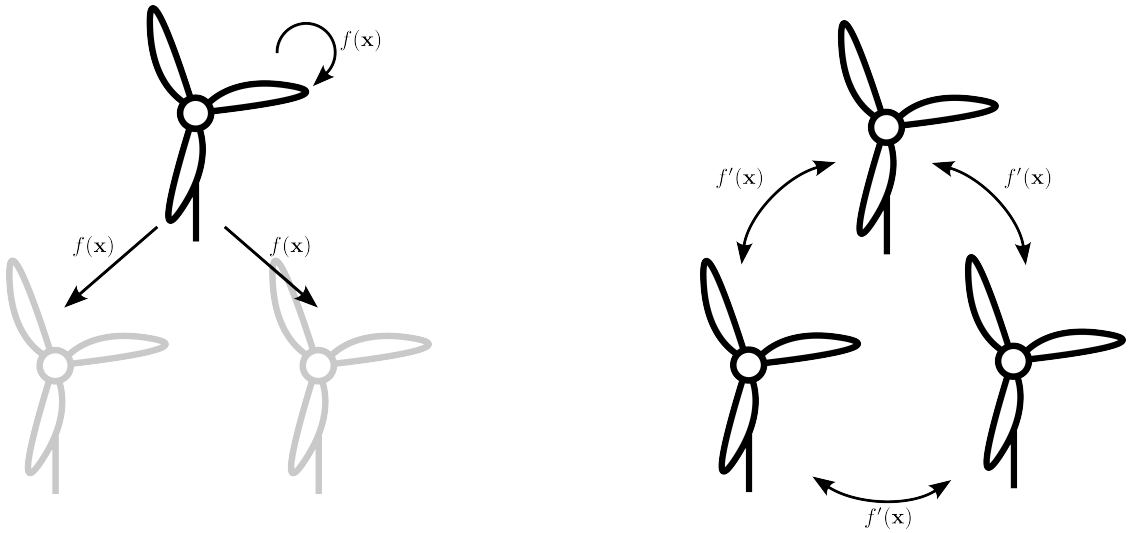
$$DEL_{LT} = \left(\frac{1}{n} \sum_{i=1}^{n} DEL_i^m \right)^{1/m} \tag{2}$$

The availability of strain measurements (and therefore, DELs) at three locations opens the door at two different methodological philosophies: fleet-leader models and population-based models (see Figure 1). In the first, a single instrumented wind turbine is used for the training of a model which then extrapolates farm-wide [21]. This hypothesis isn't that far-fetched, as the population of structures within a wind farm are tendentially homogeneous (nominally-identical machines) and the inclusion of accelerations in the training dataset is believed (and empirically confirmed [9]) to sufficiently cover the variability within the farm for most operational cases. The second philosophy, hinges on data availability on a population level. Here, we follow the grammar and discussion of [22], where models of a given feature space attempt to capture the *form* of the object of interest, *i.e.*, the *'essential'* nature of the object (let's call it the baseline behaviour, shared across the population) and the variations (deviations from this ideal essential nature) found across any real-world population, in a latent space shared by the population. In the specific case of this contribution, the implementation of a fleet-leader philosophy would lead to a latent function, f, trained on strain gauge-instrumented turbine. As for a population-based approach, it is only possible to have DEL measurements for three turbines, thus not a full population. However, one can argue that, as each of the instrumented turbines represents a design cluster (dependent on the seabed depth, which ought to be the only variation of note between turbines), a latent function, f', of this sample of the population would be sufficient to accurately represent the form and variation of the farm through a latent space shared by all turbines.

In this contribution we attempt to guide the neural network learning by including physical knowledge specific to the problem at hand, in a so-called physics-guided machine learning approach ($\Phi-$ML). Specifically, we include the Minkowski logarithmic error (MLE) introduced in [23], described by Equation 3 for $m = 5$ and $\mathbf{Y}, \hat{\mathbf{Y}}$, the measured and predicted vectors. It is based on the L^p norm and the logarithm function, and attempts to prioritise long-term DEL performance (DEL_LT), whilst maintaining ten-minute level prediction accuracy.

$$\mathcal{L}(\mathbf{Y}, \hat{\mathbf{Y}}) = \left(\sum_{i=0}^{n} |\log(y_i + 1) - \log(\hat{y}_i + 1)|^m \right)^{1/m} \tag{3}$$

The final architecture of the neural network model was thus achieved two-fold: firstly, by using a recursive feature elimination with built-in cross-validated selection of the best number of features (RFECV [24]). Secondly, using a Bayesian hyperparameter optimization algorithm [25]), with hyperparameters $h \in \mathcal{H} = \{1, \ldots, 5\}$ hidden layers, $n \in \mathcal{N} = \{32, 64, 96, \ldots, 512\}$ neurons, $a \in \mathcal{A} = \{\text{ReLU}, \text{GELU}, \text{SELU}\}$ activation



(a) Fleet-leader. (b) Population-based.

Figure 1: Different training philosophies' diagrams. (a), fleet-leader concept, wherein a neural network model (latent function f for inputs $\mathbf{x}$) is trained for a single turbine (fleet-leader) and applied to the remainder turbines. (b), population-based concept, where the latent function f' is trained based on data from all instrumented turbines (population).

function types [26], $d \in \mathcal{D} = \{0, 0.1, 0.2, 0.3\}$ dropout rate and $o \in \mathcal{O} = \{1 \cdot e^{-2}, 1 \cdot e^{-3}, 1 \cdot e^{-4}\}$ learning rate of the optimizer (Adam [27]). The tracked loss was the MLE and the monitored metrics the mean squared error and mean absolute error. The full methodological overview can be seen in Figure 2.

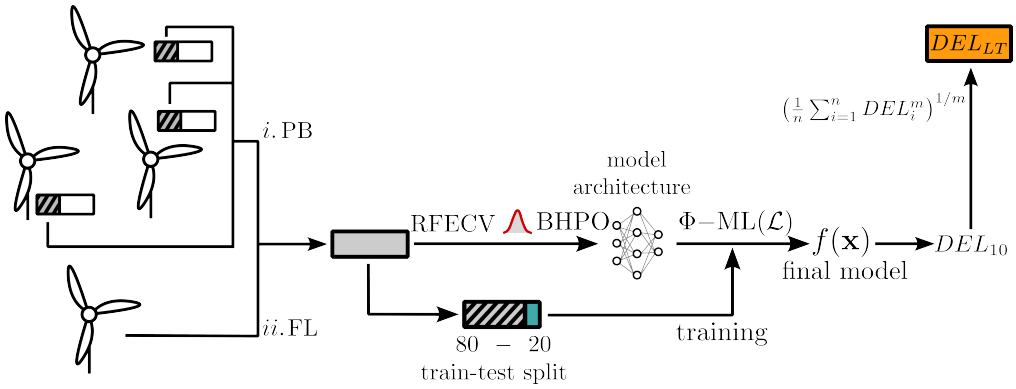


Figure 2: Overview of methodology employed, using either, i. a population-based approach (PB), where we randomly pick a third of each turbine's data, or ii. a fleet-leader approach (FL). The model architecture is attained by performing dimensionality reduction through a recursive feature estimation algorithm with cross-validation (RFECV) and through Bayesian hyperparameter optimization (BHPO). The final model (latent function, f) employs physics-guided machine learning ($\Phi-$ML) through its loss function ($\mathcal{L}$, the Minkowski logarithmic error) and estimates ten-minute DELs (DEL_{10}), further re-scaled long-term (DEL_{LT}).

RESULTS AND DISCUSSION

Following the aforementioned methodology, three fleet-leader (WT11, WT14, WT18) and a population-based model types were trained, each with five runs where the only change was the random seed value used for a 80-20 train-test split to ensure variability. In Figure 3 we can observe the long-term fore-aft DEL relative error (δ_{LT}) for each of the instrumented turbines (and the average error) based on the training dataset (turbine). In this figure we can observe several relevant differences between the models trained on

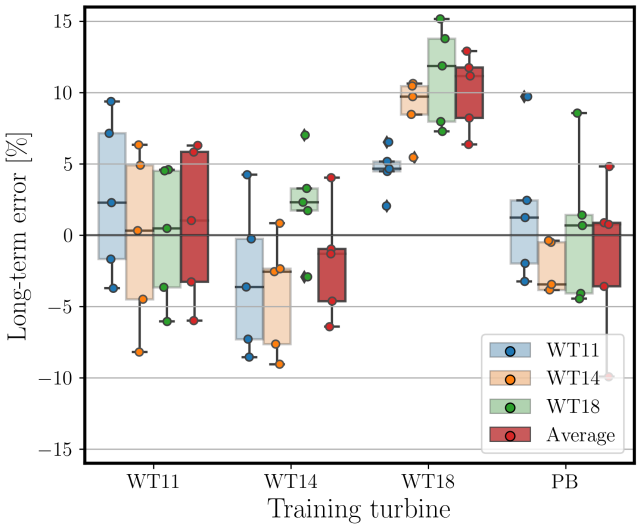
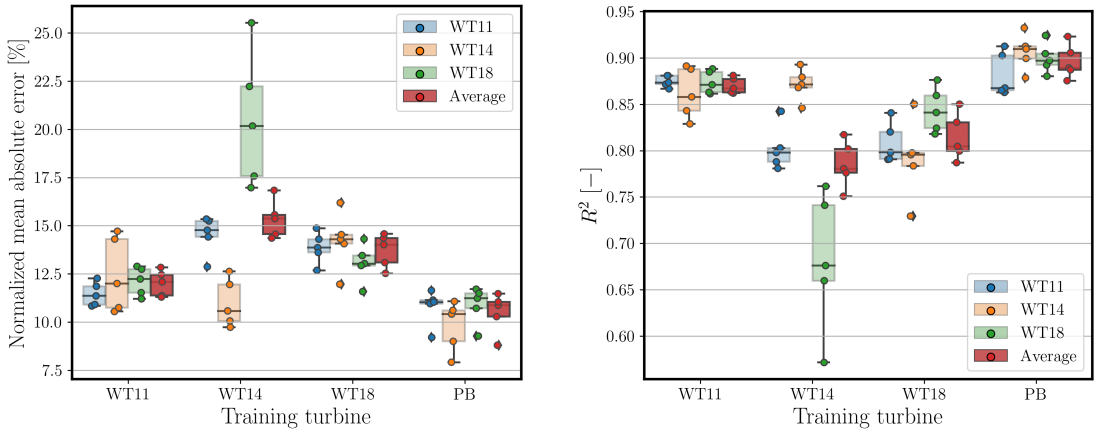


Figure 3: Box plot of long-term fore-aft DEL relative error (δ_{LT}, [%]) over three months based on the training turbine models (five runs).

different datasets. Firstly, when taking a look at the fleet-leader models (WT11, WT14, WT18), we can observe that some perform better than others, *i.e.*, models trained on WT11 have errors centered around zero with a spread of $\pm 10\%$, WT18 has a much higher average error for all turbines (centered at $+10 \pm 5\%$), while WT14 falls somewhere in the middle. This can mean that WT11 faces a higher variability of loading during its operation, which allows it to better encompass the cases found at the other locations. Additionally, the models trained using a population-based approach (combining a third of the data of each turbine) appear to be the best performing, with errors centered around zero and a smaller spread than WT11, as well as the smallest average δ_{LT}. This is to be expected, as the population-based approach combines data from the three instrumented turbines, providing a wider coverage of the operational conditions faced by the turbines. One important caveat on this and the ensuing figures is that one shouldn't conclude performance solely based on the average δ_{Lt}, but also w.r.t the individual performance for the instrumented turbines.

Albeit not being the end-goal of this contribution, one can also confer the ten-minute level model performance based on the mean absolute error, normalized w.r.t the mean DEL (NMAE) and the coefficient of determination (R^2), as seen in Figure 4.

The information transmitted by Figure 4a and Figure 4b can be very succinctly resumed as confirming the hypotheses of Figure 3: the population-based model has the lowest



(a) (b)

Figure 4: Models' ten-minute level performance based on (a) mean absolute error, normalized w.r.t the mean DEL and (b) the coefficient of determination.

NMAE and the best R^2 (above 0.9, which is rather positive), followed by WT11.
We can also take a look at the results for the side-to-side models, plotting their δ_{LT} as in Figure 3.

Here again we can repeat the previous conclusions in broad-strokes: WT11 produces

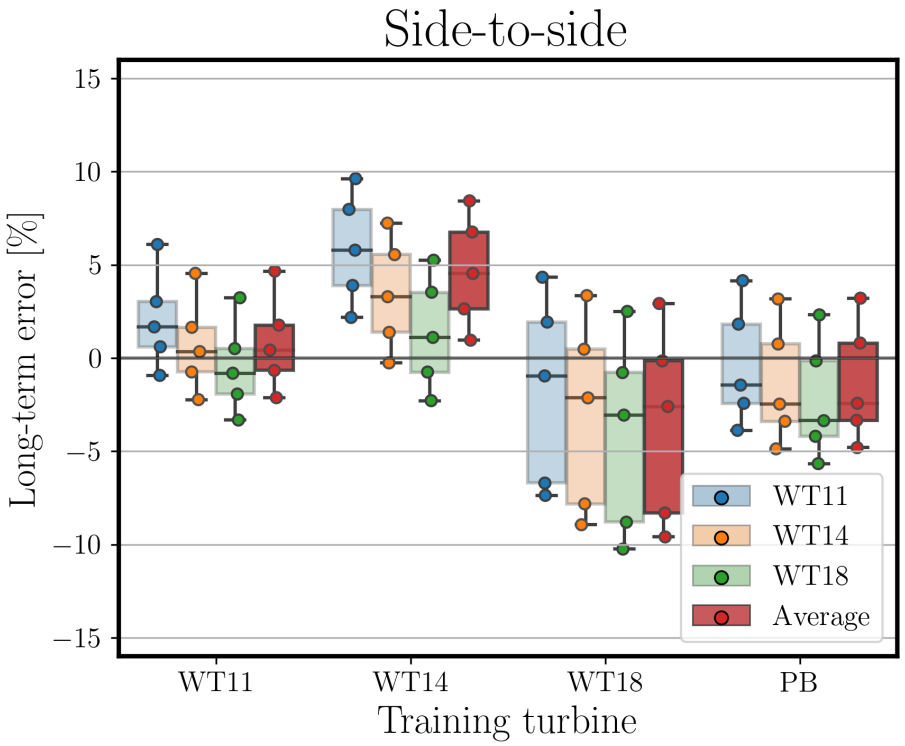


Figure 5: Box plot of long-term side-to-side DEL relative error (δ_{LT}, [%]) over three months based on the training turbine models (five runs).

better models than the other fleet-leaders, but this time around, the distinction between WT11 and the population-based models is less clear-cut. Both perform rather well, albeit WT11 has a slight positive error bias (*i.e.*, under-predicting models), whilst the inverse is true for PB. Both present a spread of $\pm 5\%$. A possible explanation for why these models are so similar is two-fold: firstly, the dynamics faced by instrumented turbines are already rather similar from the get-go; secondly, by sampling one third of the data of each turbine, the population-based approach can also loose some relevant information. But generally, we can say that for fore-aft and side-to-side, the population-based approach is better able to represent the underlying form of the farm, with there being a clear trade-off: a PB strategy doesn't allow for cross-validation in the truest sense: one can't cross-validate the model for an unseen turbine, although one can test PB models in the remainder two thirds of unseen data from the three instrumented turbines. In this contribution, the choice lies with carrying with the best-performing PB model, as it is believe to better generalize farm-wide.

Thus, we can finally plot the farm-wide normalized monthly fore-aft DEL (long-term DEL re-scaling for the month of January) for four *'representative'* turbines in Figure 6.

In this figure, one can observe how the results appear to be physically meaningful at

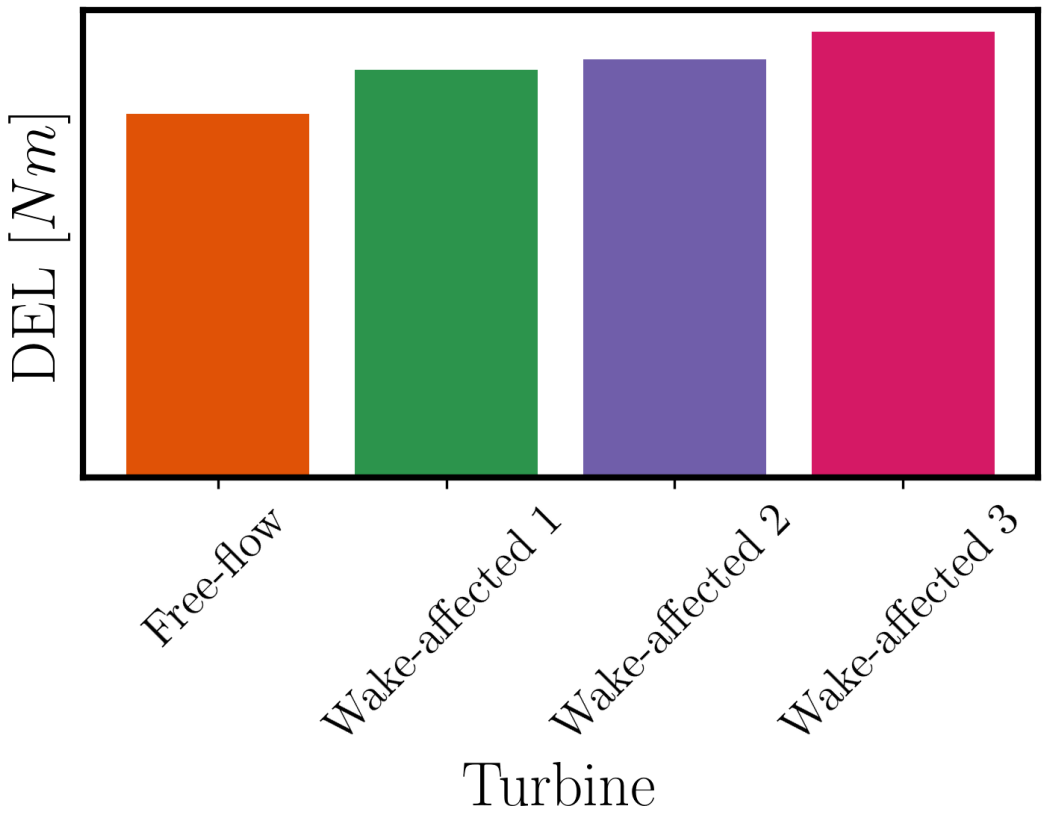


Figure 6: Farm-wide plot of normalized monthly fore-aft DEL (January) for turbines representative of free-flow and progressively wake-affected strings.

first-sight: the turbine turbine representative of the free-flow string (predominantly free-stream facing locations) presents a lower monthly DEL, which is physically coherent. Thereafter, one can also observe how the DEL increases in the wake propagation direction, with each increasing wake-affected string. Finally, this figure can furthermore be seen as the culmination of this methodology, developed first as a proof-of-concept in [1], and here presented for a real-world farm-wide accelerometer installation: a physically meaningful farm-wide monthly quantification of DEL.

This end-result is believed to be highly relevant for the industry at-large: whereas previous models could safely be assumed to accurately represent relative farm-wide behaviour, which might help to pinpoint troublesome assets, this is the first (to the best of the authors knowledge) real-world implementation of a data-driven algorithm able to accurately (within $\pm 5\%$) keep tabs on long-term DEL (and therefore, fatigue damage). As discussed in [1], a $\pm 5\%$, long-term DEL error compares favourably to current industry standards.

Translated to an industry-friendly tool, this algorithm could allow operators to track fatigue progression on their farm(s) throughout the years and serve as a springboard for further decision-making.

CONCLUDING REMARKS AND FUTURE STEPS

This contribution has offered a first view into the real-world implementation of a data-driven physics-guided machine learning model for tower fore-aft and side-to-side long-term DEL estimation based on SCADA and acceleration data. The overall methodology was explained, with a particular focus on the distinction between a fleet-leader and a population-based philosophies. In the results section it was seen how the long-term error both for fore-aft and side-to-side can be kept well below $\pm 5\%$, without loosing ten-minute level accuracy. Unprecedented in its scale, this exercise is the practical culmination of preceding proof-of-concept works, and can be seen as pointing towards a practice that will become more prevalent over the years in SHM for wind turbines: farm-wide accelerometer installation at nacelle-level in order to keep tabs of the fatigue progression for every turbine within the farm. Work that ought to be short-term presented includes the further assessment of physical meaning of the trained models by analysing outlier behaviour and performance under different operating conditions, along with the leveraging of these models' outputs for specific exercises.

ACKNOWLEDGMENTS

The authors would like to acknowledge 24SEA and Parkwind for the data collection and management and the financing provided by the ICON Supersized 4.0 project.

REFERENCES

1. de N Santos, F., P. D'Antuono, K. Robbelein, N. Noppe, W. Weijtjens, and C. Devriendt. 2023. "Long-term fatigue estimation on offshore wind turbines interface loads through loss function physics-guided learning of neural networks," *Renewable Energy*.
2. de N Santos, F., K. Robbelein, P. D'Antuono, N. Noppe, W. Weijtjens, and C. Devriendt. 2023. "Towards a Fleetwide Data-Driven Lifetime Assessment Methodology of Offshore Wind Support Structures Based on SCADA and SHM Data," in *European Workshop on Structural Health Monitoring*, Springer, pp. 123–132.
3. Sparrevik, P. 2019. "Offshore Wind Turbine Foundations State of the Art," in *From Research to Applied Geotechnics*, IOS Press, pp. 216–238.
4. García, D. and D. Tcherniak. 2019. "An experimental study on the data-driven structural health monitoring of large wind turbine blades using a single accelerometer and actuator," *Mechanical Systems and Signal Processing*, 127:102–119.
5. Wymore, M. L., J. E. Van Dam, H. Ceylan, and D. Qiao. 2015. "A survey of health monitoring systems for wind turbines," *Renewable and Sustainable Energy Reviews*, 52:976–990.
6. Worden, K., C. R. Farrar, G. Manson, and G. Park. 2007. "The fundamental axioms of structural health monitoring," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463(2082):1639–1664.
7. Iliopoulos, A., W. Weijtjens, D. Van Hemelrijck, and C. Devriendt. 2017. "Fatigue assessment of offshore wind turbines on monopile foundations using multi-band modal expansion," *Wind Energy*, 20(8):1463–1479.
8. Noppe, N., A. Iliopoulos, W. Weijtjens, and C. Devriendt. 2016. "Full load estimation of an offshore wind turbine based on SCADA and accelerometer data," in *Journal of Physics: Conference Series*, IOP Publishing, vol. 753, p. 072025.
9. de N Santos, F., N. Noppe, W. Weijtjens, and C. Devriendt. 2022. "Data-driven farm-wide fatigue estimation on jacket-foundation OWTs for multiple SHM setups," *Wind Energy Science*, 7(1):299–321.
10. Movsessian, A., M. Schedat, and T. Faber. 2021. "Feature selection techniques for modelling tower fatigue loads of a wind turbine with neural networks," *Wind Energy Science*, 6(2):539–554.
11. Avendaño-Valencia, L. D., I. Abdallah, and E. Chatzi. 2021. "Virtual fatigue diagnostics of wake-affected wind turbine via Gaussian Process Regression," *Renewable Energy*, 170:539–561.
12. Mylonas, C. and E. Chatzi. 2021. "Remaining useful life estimation for engineered systems operating under uncertainty with causal GraphNets," *Sensors*, 21(19):6325.
13. Dirlik, T. 1985. *Application of computers in fatigue analysis*, Ph.D. thesis, University of Warwick.
14. Marsh, G., C. Wignall, P. R. Thies, N. Barltrop, A. Incecik, V. Venugopal, and L. Johanning. 2016. "Review and application of Rainflow residue processing techniques for accurate fatigue damage estimation," *International Journal of Fatigue*, 82:757–765.
15. Kauzlarich, J. 1989. "The Palmgren-Miner rule derived," in *Tribology Series*, Elsevier, vol. 14, pp. 175–179.

16. Ziegler, L. and M. Muskulus. 2016. “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*, American Society of Mechanical Engineers, vol. 49972, p. V006T09A054.
17. Hübler, C., W. Weijtjens, R. Rolfes, and C. Devriendt. 2018. “Reliability analysis of fatigue damage extrapolations of wind turbines using offshore strain measurements,” *Journal of Physics: Conference Series*, 1037(3):032035.
18. Hendriks, H. and B. Bulder. 1995, “Fatigue Equivalent Load Cycle Method,” .
19. Cosack, N. 2010. *Fatigue load monitoring with standard wind turbine signals*, Ph.D. thesis, Universität Stuttgart.
20. Seidel, M., S. Voormeeren, and J.-B. van der Steen. 2016. “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):583–590.
21. Noppe, N., C. Hübler, C. Devriendt, and W. Weijtjens. 2020. “Validated extrapolation of measured damage within an offshore wind farm using instrumented fleet leaders,” in *Journal of Physics: Conference Series*, IOP Publishing, vol. 1618, p. 022005.
22. Bull, L., P. Gardner, J. Gosliga, T. Rogers, N. Dervilis, E. Cross, E. Papatheou, A. Maguire, C. Campos, and K. Worden. 2021. “Foundations of population-based SHM, Part I: Homogeneous populations and forms,” *Mechanical systems and signal processing*, 148:107141.
23. de N Santos, F., P. D’Antuono, N. Noppe, W. Weijtjens, and C. Devriendt. 2022. “Minkowski logarithmic error: A physics-informed neural network approach for wind turbine lifetime assessment,” in *Proceedings of the 30th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*.
24. Guyon, I., J. Weston, S. Barnhill, and V. Vapnik. 2002. “Gene selection for cancer classification using support vector machines,” *Machine learning*, 46:389–422.
25. Snoek, J., H. Larochelle, and R. P. Adams. 2012. “Practical bayesian optimization of machine learning algorithms,” *Advances in neural information processing systems*, 25.
26. Ramachandran, P., B. Zoph, and Q. V. Le. 2017. “Searching for activation functions,” *arXiv preprint arXiv:1710.05941*.
27. Kingma, D. P. and J. Ba. 2014. “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*.