Data-Driven Modelling of the Wake of a Pitching Porous Disk

G. Van Essche[1,2], T. Haas[1,2], J. Decuyper[1,2], T. De Troyer[1,2], M. C. Runacres[1,2]

[1] Thermo and Fluid Dynamics (FLOW), Faculty of Engineering, Vrije Universiteit Brussel (VUB), Pleinlaan 2, Brussels, 1050, Belgium
[2] Brussels Institute for Thermal-Fluid Systems and Clean Energy (BRITE), Vrije Universiteit Brussel (VUB) and Université libre de Bruxelles (ULB), Brussels, 1050, Belgium

Keywords: Data-Driven Modelling, Unsteady and High Dimensional Aerodynamics, Porous Disk.

Abstract. Porous disk systems are widely used in wind tunnel experiments as an initial step toward modelling Floating Offshore Wind Turbines (FOWTs). While FOWTs have great potential in wind energy, their unsteady aerodynamics pose challenges for performance, lifespan, and wind farm optimisation. Their inherent nonlinearity, unsteadiness, and high dimensionality make modelling computationally expensive. Data-driven approaches offer a promising alternative by reducing computational costs while maintaining accuracy. Evaluating these methods on simpler surrogate systems helps address key challenges before scaling up to full FOWT models. In this study, CFD simulations of a pitching actuator disk are conducted to mimic the rotor aerodynamics of an IEA 15MW offshore wind turbine. Ten velocity time series are collected from the wake of a pitching porous disk under single-sine input signals of varying amplitudes and frequencies. An encoder-propagator-decoder framework, combining a convolutional neural network (CNN) and a long short-term memory (LSTM) network, is tested on five of these cases. The CNN encoder achieves high accuracy, with a reconstruction error of approximately 1%. The propagator's accuracy varies across cases, ranging from a few percent to around 20%. These findings demonstrate the potential of data-driven techniques for reconstructing high-dimensional velocity fields and integrating them into wake control strategies.

1. Introduction

In this work, we investigate the response of a pitching porous disk as a surrogate system for Floating Offshore Wind Turbines (FOWTs). FOWTs represent a significant advancement in wind engineering technology, particularly as available land and shallow-water sites become increasingly scarce [1]. However, the inherent motion of these floating platforms amplifies unsteady aerodynamic effects at the rotor compared to fixed-bottom turbines, impacting energy production, structural lifespan, and optimal wind farm layout. Characterising FOWT wakes is crucial not only for wake control strategies based on platform motion [2] but also for wind farm layouts optimisation. These wake systems are inherently unsteady, non-linear, and high-dimensional, adressing significant challenges for high-fidelity first-principle models, which are computationally too expensive for practical integration into control frameworks. A promising alternative is to leverage data-driven models trained on generated data, making them more suitable for implementation within control strategies. A first step toward full-scale data-driven modelling, is to evaluate the potential of these techniques on simpler systems that retain key challenges such as nonlinearity, unsteadiness, and high dimensionality. The pitching porous disk system encapsulates these complexities and is widely used in wind tunnel experiments, making it an ideal starting point for data-driven modelling.

Previous studies have demonstrated the potential of data-driven approaches. For instance, Damiola et al. [3] successfully employed a state-space neural network (SS-NN) [4] to capture the relationship between the time-varying angle of attack of a pitching NACA 0018 aerofoil and its lift coefficient using CFD data. Similarly, Decuyper et al. [5] applied a discrete-time polynomial non-linear state-space model to CFD data from a cylinder oscillating in a cross-flow, accurately capturing the

interaction between the cylinder's motion and the resulting fluid forces. Additionally, Lee and You [6] predicted flow fields by employing deep learning networks with encoder-propagator-decoder architectures based on historical flow field data, achieving results in close agreement with numerical simulations.

The primary objective is to assess the potential of data-driven methods in developing a framework capable of generating stable and accurate multi-step-ahead predictions of the near wake from a given input signal. This framework is designed ensuring readily integration into model predictive control (MPC) for floating offshore wind turbines (FOWTs). To achieve this, an encoder-propagator-decoder architecture is implemented. A convolutional neural network (CNN) [7] is employed to encode high-dimensional data into a reduced spatial representation. Subsequently, a long short-term memory (LSTM) [8] network propagates the data forward in time within the reduced space. Finally, the CNN is used again to decode the simulated data back into the full-dimensional space. These models are trained and tested on an unsteady dataset generated from a pitching porous disk, designed to replicate the rotor aerodynamics of an IEA 15MW offshore reference wind turbine [9]. The models take the time series of the disk's pitching angle as input, while their outputs correspond to velocity magnitudes measured at one hundred spatial locations within the near wake. Training, validation and testing are conducted for different cases corresponding to different combination of amplitude and frequency of the input signal.

2. Data

The wind flow around the floating offshore wind turbine is emulated by performing unsteady Reynolds-Averaged Navier-Stokes (uRANS) simulations of a porous disk using OpenFOAM v2406. The porous disk has a diameter D=240 m and a thickness of 5 m and the porosity ratio of the disk area is 33%. The disk diameter corresponds to the diameter of the IEA 15MW offshore reference wind turbine [9].

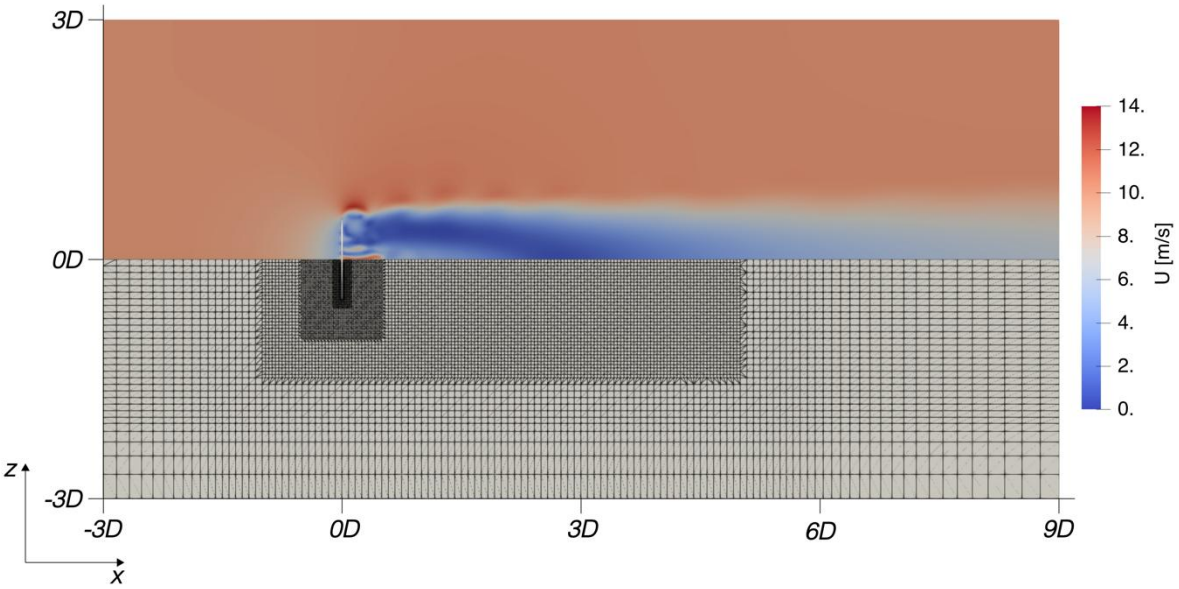

Figure 1: Side view of (top) the instantaneous velocity magnitude at t = 800 s for A = 2 deg and f = 0.05 Hz and (bottom) the computational mesh.

The computational domain spans over 12D in the streamwise direction, 6D in the spanwise direction, and 6D in the vertical direction. The disk is located in the middle of the crossflow plane, respectively 3D downstream of the inflow and 9D upstream of the outflow boundaries. The blockage ratio of the total disk area equals 2.18%. The domain is discretised into about 1.8 million cells using local surface and region refinement. The first refinement region extends from 1D upstream to 5D downstream of the disk

in the streamwise direction and spans over $3D$ in both the spanwise and vertical directions in order to resolve the near-wake region with a diameter-to-cell ratio D/Δ=24. The mesh is further refined in the disk region, expanding $1.5D$ away from the disk in all three directions, with three successive refinement levels resulting in a diameter-to-cell ratio D/Δ=192 in the vicinity of the disk. The computational domain and the mesh are illustrated on Figure 1.

The Reynolds stresses are modelled using the $k-\varepsilon$ turbulence model. Time discretisation is achieved using a first order implicit Euler scheme. Spatial discretisation of gradient and Laplacian terms is done using second order central schemes while divergence terms employ second and first order upwind schemes for velocity and turbulence quantities respectively.

A fixed velocity inlet condition with a uniform inflow speed U=10 m/s is combined with a zero-pressure outlet condition and slip conditions at the domain sides. The turbulence intensity at the inflow is set to 0.5%. A moving wall velocity condition and wall models for turbulence quantities are applied to the boundaries of the porous disk. An oscillating pitching motion of amplitude A and frequency f, originating from 150 m below the disk center, is imposed to the disk in combination with mesh morphing. Starting from steady-state solutions of the disk at standstill, the flow equations are solved for a time horizon T=800 s using the PISO algorithm with two pressure correction steps. The simulation time step is set to Δt=0.01 s to ensure CFL<1. Simulations are performed on the Tier-2 HPC infrastructure of the VUB using 32 processors of AMD EPYC 9384X nodes, requiring a total of 96 node-hours of compute time for the generation of the dataset.

Velocity measurements are collected at 100 grid locations, uniformly spaced between $1D$ and $2.5D$ along the x-direction and between $-\frac{1}{2}D$ and $\frac{1}{2}D$ along the z-direction. The locations are evenly spaced at 40 m intervals. Data is sampled at a frequency of 20 Hz. All simulations available and their corresponding case number are shown in Table 1.

Table 1. Simulations cases and the corresponding amplitude A [°] and frequency f [Hz] of the oscillating pitching motion.

Case #	1	2	3	4	5	6	7	8	9	10
$[A; f]$	[2; 0.05]	[2; 0.10]	[2; 0.15]	[2; 0.20]	[4; 0.05]	[4; 0.10]	[4; 0.15]	[4; 0.20]	[2.5; 0.12]	[3.5; 0.12]

3. Models Set-up

The training setup of the model is illustrated in Figure 2. First, a convolutional neural network (CNN) is trained to reduce the 100 spatial dimensions into a four-dimensional latent space. Next, a long short-term memory (LSTM) network is trained using the pitching input and the reduced velocity field. For this step, the data is organised into batches with predefined sequence lengths, determining how many previous time steps the LSTM model considers to predict the next time step.

The CNN autoencoder (AE) consists of two 2D convolutional layers with ReLU activation functions. The first layer reduces the spatial dimensions from 10×10 to 6×6, and the second further compresses it to 2×2, resulting in a four-dimensional latent space with a kernel size of 5. The CNN model contains 6,833 parameters and is trained using the Adam optimizer [10].

The LSTM propagator comprises a single layer with 64 neurons and a tanh activation function. The data is structured into batches of sequences with a length of 600 and a stride of 10, corresponding to a time window of 30 seconds. The LSTM network consists of 17,412 parameters and is also trained using the Adam optimizer [10].

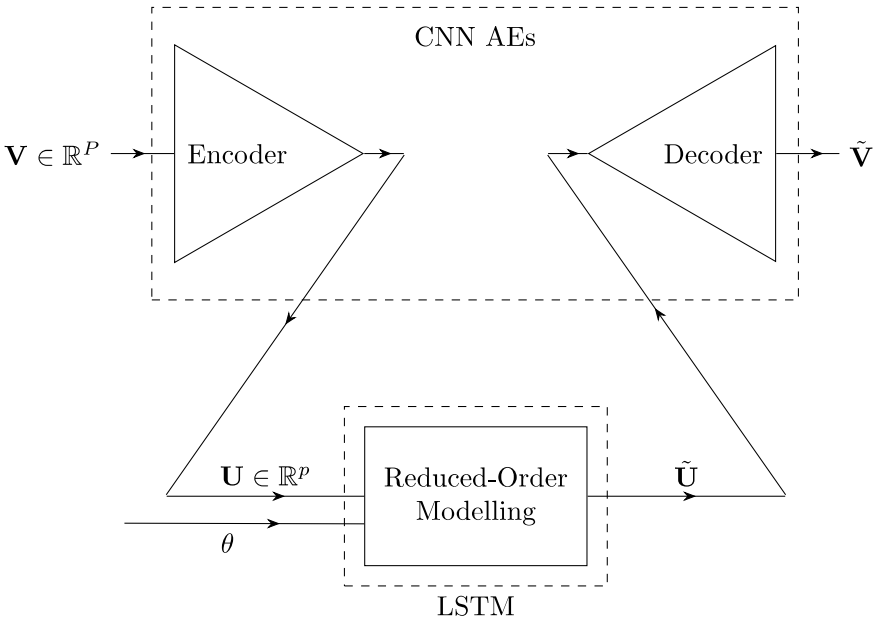

Figure 2: Encoder-propagator-decoder architecture of the model with **V** the high dimensional velocity field of dimension $P = 100$, U the latent space of dimension $p = 4$ and θ the pitching motion.

4. Results

The full length of a time series of a velocity measured at one of the grid locations in the wake of the CFD simulation is shown on Figure 3 below.

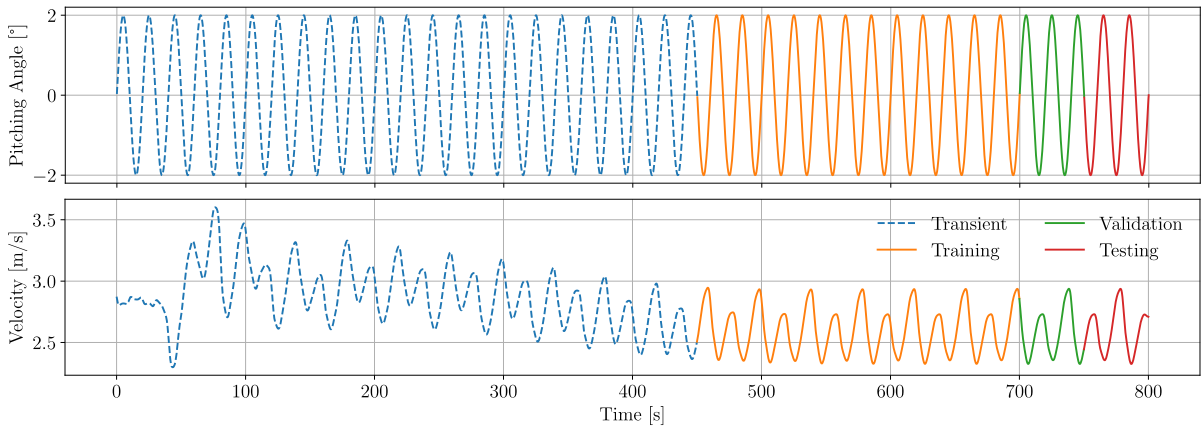

Figure 3: Time series of pitching angle (up) and of velocity measured at position (360, 160) m in the wake for Case 1 (bottom).

As illustrated in the figure above, transient phenomena persist up to approximately $t = 450$ s, as the transient effects need to dissipate in the wake. This portion of the time series is disregarded. The remaining data is then divided into training, validation, and testing sets before training the data-driven models. The training set contains 5,000 samples, while the validation and testing sets each contain 1,000 samples.

The method's performance is evaluated on four cases using this data split. For all cases, the data was normalised, and the propagator model was trained multiple times. The most accurate model was selected to mitigate the influence of suboptimal random weight initialisation. The results for Cases 1, 4, 5, and 9 are presented in Table 2.

Table 2. Results of the data-driven method for Case 1, 4, 5, and 9. All errors are NRMSE.

		Training	Validation	Testing
Case 1 [2; 0.05]	**Reconstruction Error**	0.87 %	0.78 %	0.85 %
	ROM Error	8.22 %	8.29 %	8.22 %
Case 4 [2; 0.20]	**Reconstruction Error**	3.30 %	0.61 %	2.05 %
	ROM Error	25.69 %	30.69 %	19.92 %
Case 5 [4; 0.05]	**Reconstruction Error**	0.52 %	0.43 %	0.43 %
	ROM Error	2.13 %	1.42 %	2.79 %
Case 9 [3.5; 0.12]	**Reconstruction Error**	1.93 %	0.49 %	0.57 %
	ROM Error	7.36 %	7.35 %	7.37 %

The results demonstrate that the CNN AE effectively compresses and reconstructs the velocity field, with reconstruction errors consistently around 1. % across all cases. This validates the ability of the CNN AE to accurately encode and decode the 100 spatial velocity dimensions into a four-dimensional latent space. However, the results suggest a trend where the reconstruction accuracy slightly deteriorates as the frequency of the pitching motion increases. This is particularly evident in Case 4, where the highest frequency (0.20) corresponds to a higher reconstruction error (3.30 % in training and 2.05 % in testing), compared to lower-frequency cases such as Case 5 (0.43 %).

The accuracy of the reduced-order model (ROM) propagator exhibits more variability depending on the specific case. A key observation is that higher pitching frequencies tend to introduce greater errors in the ROM predictions. Case 4, with the highest frequency (0.20), has the largest ROM errors, reaching 30.69 % in validation. This aligns with the expectation that higher-frequency motions introduce more complex transient dynamics, making accurate propagation more challenging. However, the influence of pitching amplitude on ROM accuracy is less straightforward. For example, despite Case 5 having a larger amplitude (4) than Case 1 (2), its ROM errors are significantly lower (maximum 2.79 %) compared to Case 1 (maximum 8.29 %). This suggests that while frequency strongly affects model accuracy, amplitude alone is not a clear predictor of performance. Other factors, such as the interplay between frequency and amplitude or non-linear wake interactions, may contribute to these variations.

Overall, these results highlight the strengths of the CNN AE in capturing spatial features and the challenges faced by the LSTM-based propagator, particularly in high-frequency cases. Further investigation into optimising the propagator's robustness to high-frequency inputs and exploring alternative architectures may help improve predictive performance in complex flow conditions.

5. Conclusion

Although the proposed work demonstrates the potential of data-driven methods to accurately simulate high-dimensional velocity wakes, several improvements are necessary before these models can be implemented in a control framework. Further research is required in both the CFD and data-driven frameworks before advancing toward such applications.

The introduced CFD framework is still in its early stages, and its accuracy will be refined in future work. The disk porosity will be adjusted to match the thrust coefficient of the IEA 15MW offshore

reference wind turbine, and a grid convergence study will be conducted to verify the computational framework's reliability.

The data-driven framework will be further evaluated using broadband input signals. Additionally, training, validation, and testing procedures will be enhanced by incorporating multiple concatenated experiments within each dataset to improve model generalisation. While this research presents significant challenges and requires further development, the authors are confident that these frameworks will soon be capable of providing a robust foundation for full-scale FOWT applications.

6. Acknowledgments

This work was funded by the FWO fellowship under project number 1S53825N.

7. References

[1] P. Veers, K. Dykes, S. Basu, A. Bianchini, A. Clifton, P. Green, H. Holttinen, L. Kitzing, B. Kosović, J. K. Lundquist, J. Meyers, M. O'Malley, W. J. Shaw, B. Straw 'Grand Challenges: wind energy research needs for a global energy transition.', Wind Energy Science, https://doi.org/10.5194/wes-7-2491-2022
[2] E. M. Nanos, C. L. Bottasso, S. Tamaro, D.I. Manolas, V. A. Riziotis 'Vertical wake deflection for floating wind turbines by differential ballast control.', Wind Energy Science, https://doi.org/10.5194/wes-7-1641-2022
[3] L. Damiola, J. Decuyper, M. C. Runacres, T. De Troyer 'Modelling the unsteady lift of a pitching NACA 0018 aerofoil using state-space neural networks.', Journal of Fluid Mechanics, https://doi.org/10.1017/jfm.2024.148
[4] J. A. Suykens, B. L. De Moor, J. Vandewalle 'Nonlinear system identification using neural state space models, applicable to robust control design.', International Journal of Control, https://doi.org/10.1080/00207179508921536
[5] J. Decuyper, T. De Troyer, M. C. Runacres, K. Tiels, J. Schoukens 'Nonlinear state-space modelling of the kinematics of an oscillating circular cylinder in a fluid flow.', Mechanical Systems and Signal Processing, https://doi.org/10.1016/j.ymssp.2017.04.048
[6] S. Lee, D. You 'Data-driven prediction of unsteady flow over a circular cylinder using deep learning.' Journal of Fluid Mechanics, https://doi.org/10.1017/jfm.2019.700
[7] N. Ketkar, J. Moolayil, 'Convolutional Neural Networks', book chapter pp 197–242, Deep Learning with Python. Apress, Berkeley, 2021; DOI: https://doi.org/10.1007/978-1-4842-5364-9_6
[8] S. Hochreiter. Long short-term memory. Neural Computation MIT-Press, 10.1162/neco.1997.9.8.1735
[9] E. Gaertner, 'Definition of the IEA 15-megawatt offshore reference wind', National Renewable Energy Laboratory, 2020.
[10] D. P. Kingma, J. Ba, 'Adam: A method for stochastic optimization', arXiv preprint arXiv:1412.6980, https://doi.org/10.48550/arXiv.1412.6980