Jornadas de Automática

Data-Driven Emulation of Wind Turbine Pitch Control with an RBF Network Under Varying Turbulence Conditions

Lei Ren [a], Matilde Santos [b,*], J. Enrique Sierra-García [c], Xin Cai[a]

[a] College of Mechanics and Engineering Science, Hohai University, 211100, Nanjing, China.
[b] Institute of Knowledge Technology, Complutense University of Madrid, 28040, Madrid, Spain.
[c] Systems Engineering and Automation, University of Burgos, 09004, Burgos, Spain.

To cite this article: Ren, L., Santos, M., Sierra-García, J.E., Cai X.. 2026. Data-Driven Emulation of Wind Turbine Pitch Control with an RBF Network Under Varying Turbulence Conditions. Jornadas de Automática, 47. https://doi.org/10.17979/ja-cea.2026.47.13718

Abstract

Above rated wind speed, collective pitch control (CPC) is essential for wind turbines to regulate generator speed and power. This paper proposes a data-driven emulation framework that reproduces the input-output behaviour of the reference software OpenFAST CPC controller using a radial basis function (RBF) neural network with a sliding time-window feature extraction. Multi-scale statistical features—derived from generator speed error, power error, pitch angle, wind speed, and generator speed over a 30-s window—are fed into an RBF network to predict the next-step collective pitch angle. The model is trained on OpenFAST simulation data under three different load cases: DLC1.1, DLC1.3 and DLC1.6 data sets. Two training configurations (Single-DLC and Multi-DLC) are compared. Closed-loop validation under DLC1.1 and DLC1.3 datasets shows that the RBF emulator achieves high fidelity ($R^2 > 0.99$, RMSE $< 0.5°$) and closely tracks the baseline PI controller with negligible performance degradation. The proposed surrogate controller provides a lightweight, accurate alternative for scenarios where the original regulator is inaccessible or fast prototyping is required.

Keywords: Wind turbine; Intelligent control; Collective pitch control; Radial basis function network; Data-driven emulation.

1. Introduction

When wind turbines operate above rated wind speed, active pitch control is essential to limit power capture and reduce structural loads (Muñoz et al., 2024). Collective pitch control (CPC), which adjusts all three blade angles uniformly, remains the most widely adopted strategy in both industry and research due to its simplicity and effectiveness (Gambier, 2021). In the open-source simulation tool OpenFAST, CPC is implemented as a gain-scheduled proportional–integral (PI) regulator, serving as the baseline for numerous studies on load analysis, floating wind turbines, and pitch actuation dynamics (Bai et al., 2025; Jeong et al., 2025).

Considerable research has focused on improving CPC performance beyond the baseline proportional-integral PI gain scheduling controller, including reinforcement learning-based parameter optimization (Espinoza et al., 2025), fuzzy logic enhanced PI (Serrano et al., 2022), and comparisons between classical PID controllers and machine-learning approaches (Kane, 2020; Sierra-García & Santos, 2021; Velino et al., 2022). For large-scale and floating turbines, individual pitch control (IPC) has also been developed to mitigate asymmetric loads, often combined with model predictive control, H∞, or LIDAR feedforward (Routray et al., 2023; Russell et al., 2024; Zheng et al., 2025). The common theme across these efforts is the design of new control laws.

In contrast, this paper focuses on a related but relatively underexplored problem: the data-driven emulation of an existing CPC controller. A surrogate that captures the CPC input-output dynamics offers a practical alternative when the native controller cannot be directly accessed—for instance, when integrating wind turbine control logic into hardware-in-the-loop testbeds, third-party simulation tools, or platforms where modifying the original OpenFAST source code is not feasible.

The proposed emulation framework is detailed in Section 2, which covers the overall architecture, data generation and feature extraction, RBF network training, and online closed-loop implementation. Section 3 presents simulation results, including offline validation metrics and closed-loop

*Correspondence Author: msantos@ucm.es

dynamic responses under load cases DLC1.1 and DLC1.6 data sets. Section 4 concludes the paper and discusses future work.

2. Design of the Sliding-Window RBF Emulator

2.1. Overall Architecture

The approach comprises offline training and online deployment (Fig. 1). In the offline phase, OpenFAST simulations under DLC1.1 (*IEC 61400-1*, n.d.) generate time-series data. A 30 s sliding window is used to extract multi-scale statistical features, while the target is the next-step collective pitch angle (mean of the three blades). A radial basis function (RBF) neural network is then trained, with centers determined via k-means clustering algorithm and output weights computed by ridge regression. In the online phase, real-time sensor data are processed using the same feature extraction scheme, and the trained RBF network predicts the pitch command to drive the actuator in closed loop. No online retraining is performed.

To investigate the influence of training data diversity, two training configurations are considered: Single-DLC, which uses only DLC1.1 data, and Multi-DLC, which incorporates DLC1.1, DLC1.3, and DLC1.6. Both configurations employ the same network architecture and hyperparameters.

2.2. Data Generation, Preprocessing, and Feature Extraction

Simulations use the IEA-15-240-RWT reference turbine (Allen et al., 2020) in OpenFAST. For the Single-DLC configuration, only DLC1.1 is used; for the Multi-DLC configuration, three design load cases—DLC1.1, DLC1.3, and DLC1.6—are combined to cover normal turbulence, extreme turbulence, and extreme operating gust conditions. According to IEC 61400-1, DLC1.1 corresponds to the Normal Turbulence Model (NTM) and is used for fatigue analysis and extreme load extrapolation with a 50-year return period. DLC1.3 employs the Extreme Turbulence Model (ETM) for direct ultimate limit state assessment, with loads not less than those extrapolated from DLC1.1. DLC1.6 also uses NTM but is specifically applied to combined wind, wave, and current loads under severe sea state conditions.

All runs employ a sampling time of $\Delta t = 0.025$s and cover above-rated wind conditions where CPC is active. The first 100 s are discarded to remove start-up transients. Recorded variables include generator speed ω_{gen}, generator power P_{gen}, wind speed v, and blade pitch angles β_1, β_2, β_3. The collective pitch is defined as $\beta_{\mathrm{CPC}} = (\beta_1+\beta_2+\beta_3)/3$.

A multi-scale sliding window (L = 30s, stride 0.1 s) extracts features from five quantities: generator speed error error $e_\omega = \omega_{rated} - \omega_{gen}$, power error $e_P = P_{rated} - P_{gen}$, β_{CPC}, v, and ω_{gen}. For each, statistical features (current value, mean, slope, integral, standard deviation) are computed over tailored sub-windows (Table 1), right-aligned at time t_k. The target is β_{CPC} at $t_k + 0.1$ s. This yields a 14-dimensional feature vector $\mathbf{X}(t)$, capturing different system time scales.

All DLC data are merged and filtered to remove invalid samples or pitch values outside [0°, 30°]. The dataset is split by simulation into training (66.7%), validation (16.6%), and test (16.7%) sets to avoid temporal leakage, with standardization based on training data.

2.3. RBF Network Training for CPC Emulation

A single-output RBF network is employed. The hidden layer contains K = 500 Gaussian units with a common spread σ = 5.0. Centers $c_j \in \mathbf{R}^{14}$ are obtained via k-means clustering on 150,000 randomly selected training samples. For an input $\mathbf{X}$, the activation $\phi_j(\boldsymbol{X})$ of the j-th Gaussian hidden unit is:

$$\phi_j(\mathbf{X}) = \exp\left(-\frac{\left\|\mathrm{X} - c_j\right\|^2}{2\sigma^2}\right) \tag{1}$$

The design matrix $\mathbf{\Phi}$ includes a bias column. Output weights $\boldsymbol{W}$ are solved by ridge regression:

$$\boldsymbol{W} = (\mathbf{\Phi}^T\mathbf{\Phi} + \lambda I)^{-1}\mathbf{\Phi}^T y \tag{2}$$

with $\lambda = 0.1$. Where $y \in \boldsymbol{R}^N$ is the vector of normalized target pitch angles. The predicted pitch angle is:

$$\hat{\beta} = \sum_{j=1}^{K} w_j \phi_j(\mathbf{X}) + w_{bias} \tag{3}$$

followed by denormalization. Hyperparameters (K, σ) were tuned on the validation set. The hyperparameters were selected via a grid search over K, σ, and λ on the validation set, with the chosen values providing a favorable balance between accuracy and training stability.

2.4. Online Closed-Loop Implementation

In closed-loop operation (Fig. 2), a circular buffer stores the previous 30 s of sensor data. The dashed block represents the temporal RBF-based CPC controller, comprising multi-scale feature extraction and the trained RBF network. Every 0.1 s, the 14 features are updated and input to the RBF, which outputs the pitch command $\beta_{(\mathrm{ref})}$. This command is then filtered through a first-order actuator model with rate and saturation constraints before being applied to the OpenFAST turbine model.

3. Simulation Results and Discussion

3.1 Training and Validation Performance

The k-means clustering procedure converged in fewer than 300 iterations per replicate. Output weights were computed via ridge regression in a single step, avoiding the iterative optimization required by many alternative data-driven models. For instance, in our preliminary model-selection tests, a recurrent Long Short-Term Memory (LSTM) network trained on the same dataset required over 20 minutes to reach a comparable accuracy level. The Single-DLC configuration required approximately 50 s of wall-clock time, while the Multi-DLC configuration took approximately 130 s due to the larger training set.

Normalized RMSE values for the Single-DLC configuration were 0.0764 (training) and 0.0746 (validation), indicating no substantial overfitting.

Denormalized validation metrics are summarized in Table 2. The emulator achieves an R^2 of 0.9953 and a Pearson correlation of 0.9977, while the Multi-DLC emulator attains an R^2 of 0.9932 and a correlation of 0.9966. Both configurations capture over 99 % of the variance in the pitch command. The slightly higher RMSE of the Multi-DLC configuration (0.494° vs. 0.384°) is attributable to the increased data diversity introduced by load cases DLC1.3 and DLC1.6, which include more extreme wind conditions. The differential RMSE values remain low in both cases (0.080° and 0.090°), confirming smooth temporal tracking without spurious high-frequency fluctuations.

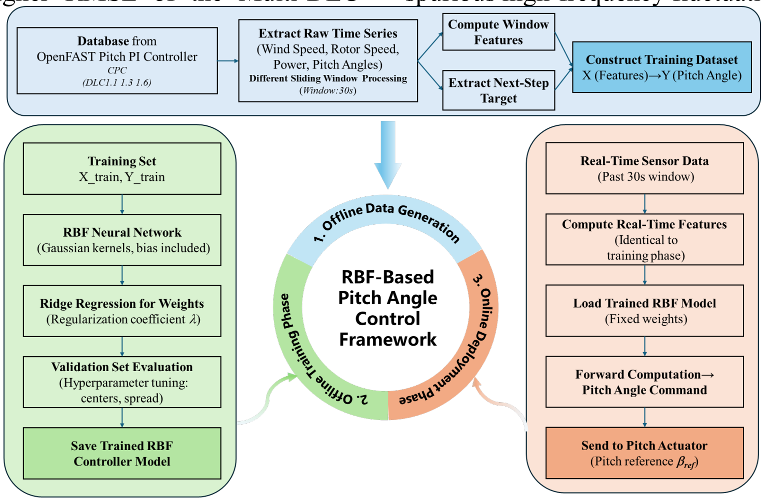

Figure 1. Overall framework of the proposed approach.

Table 1: Multi-scale feature configuration

Basic quantity	**Statistics (window length in seconds)**
Generator speed error	current (0 s), mean (30 s), slope (1.5 s), integral (15 s), standard deviation (6s)
Power error	current (0 s), integral (10 s)
CPC signal	mean(30s), slope (0.5s)
Wind speed	current (0 s), standard deviation (30 s), slope (2 s)
Generator speed	current (0 s), standard deviation (6 s)

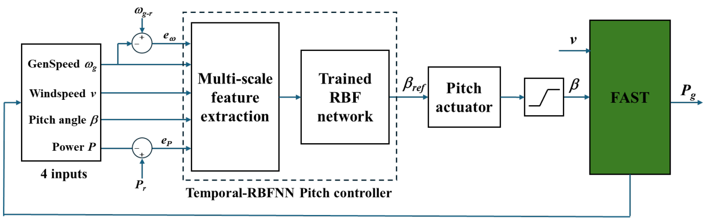

Figure 2: Closed-loop control scheme of the online deployment phase.

Table 2 Performance metrics of the RBF emulator on the validation set

Configuration	RMSE	MAE	R^2	ρ	Differential RMSE
Single-DLC	0.384°	0.302°	0.9953	0.9977	0.080°
Multi-DLC	0.494°	0.384°	0.9932	0.9966	0.090°

Figure 3 displays the scatter plots of predicted versus actual pitch angles on the test set for both configurations. In both cases, the points are tightly clustered around the diagonal, with no systematic bias.

The Single-DLC configuration covers a pitch range of approximately 2.8° to 24.5°, achieving a test-set RMSE of 0.475°, MAE of 0.325°, R^2 of 0.9926. The Multi-DLC configuration spans a slightly wider range (approximately 0° to 25°) due to the inclusion of extreme wind conditions, and yields a comparable test-set RMSE of 0.483°, MAE of 0.371°, and R^2 of 0.9925. The differential RMSE remains low in both cases (0.080° and 0.089°), and the negligible gap between validation and test RMSE confirms that neither model overfits to the particular wind profiles used during training. Notably, the Multi-DLC configuration achieves essentially the same test-set R^2 as the Single-DLC configuration despite being evaluated on a more diverse dataset that includes extreme turbulence and extreme operating gust conditions.

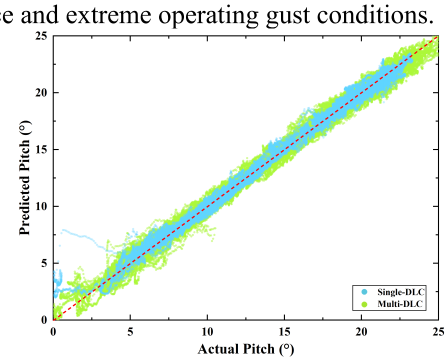

Figure 3. Scatter plot of RBF-predicted versus actual collective pitch angle on the test set (The red dashed line indicates perfect prediction.)

3.2 Closed-Loop Dynamic Validation

To assess the viability of the proposed RBF emulator as a functional substitute for the OpenFAST collective pitch controller, closed-loop simulations were conducted under a turbulent wind profile representative of DLC1.1 (mean wind speed 22 m/s, turbulence intensity 2.5%). A 700 s simulation was performed, with the first 100 s discarded to exclude start-up transients. The remaining 600 s of data, sampled at 40 Hz, provide a statistically robust basis for comparing the baseline gain-scheduled PI regulator against both the Single-DLC and Multi-DLC configurations of the RBF surrogate.

Figure 4 overlays the time-domain responses of the three controllers for hub-height wind speed, collective pitch angle, generator speed, and generator power. The wind speed profile (Figure 4a) contains multiple steep ramps and gusts exceeding 20 m/s, exercising the controllers across a broad operational envelope. As shown in Figure 4b, both RBF configurations closely track the baseline pitch trajectory. The Single-DLC model follows the reference with minor deviations on the order of 0.2°–0.5° during the most abrupt transients, and its output is slightly smoother than the baseline, as confirmed by the standard deviation of 1.447° versus 1.458° for the baseline (see Table 3).

The Multi-DLC configuration tracks the mean trend equally well, but its pitch signal exhibits small, high-frequency oscillations (roughly ±0.1–0.2°) that are absent in the Single-DLC curve. These oscillations are a consequence of the increased diversity of the Multi-DLC training set, which includes extreme turbulence and gust events from DLC1.3 and DLC1.6. The RBF network, while successfully capturing the global control mapping, becomes slightly more sensitive to rapid input variations when exposed to such data, producing minor high-frequency components in its output.

Importantly, these pitch oscillations do not propagate into the generator speed or power responses. As shown in Figure 4c and 4d, the rotor speed and power trajectories of the Multi-DLC configuration remain tightly clustered around those of the baseline and Single-DLC models, with no visible increase in fluctuation. This is because the large rotor inertia and drive-train dynamics act as a low-pass filter, effectively attenuating the high-frequency content of the pitch command before it influences the turbine's electromechanical response.

These findings are supported by the statistical summary in Table 3. Across all three controllers, the mean pitch angle, generator speed, and power differ by less than 0.5%, confirming that the RBF emulators preserve the steady-state operating point of the baseline PI regulator. The pitch standard deviations of the Single-DLC (1.447°) and Multi-DLC (1.501°) configurations are close to the baseline value (1.458°), consistent with the tight trajectory tracking observed in Figure 4b.

Despite the high-frequency oscillations noted earlier for the Multi-DLC model, its statistical dispersion remains well within the baseline bounds, and the generator speed and power statistics show no measurable impact from these pitch irregularities. The minimum and maximum values across all variables also remain within the baseline bounds, with no spurious extreme excursions, confirming stable closed-loop operation throughout the 600-second simulation.

In summary, the closed-loop validation under DLC1.1 demonstrates that both the Single-DLC and Multi-DLC configurations faithfully replicate the dynamic behavior of the native OpenFAST CPC controller. The Single-DLC model provides a smooth and accurate emulation, while the Multi-DLC model achieves the same level of accuracy in terms of mean regulation and overall dispersion, at the cost of minor high-frequency pitch oscillations that are inconsequential for the turbine's electromechanical response. These results confirm that the proposed RBF surrogate, even when trained on a single design load case, already offers a high-fidelity alternative to the original controller. The addition of multiple load cases during training does not degrade normal-operation performance and may extend the emulator's robustness to more challenging wind conditions. Nevertheless, a dedicated fatigue assessment of these oscillations is recommended for long-term maintenance considerations, and will be addressed in future studies.

To evaluate whether multi-DLC training improves robustness under more severe wind conditions, the Multi-DLC configuration was tested in closed-loop under a DLC1.3 turbulent wind profile (mean 22 m/s, turbulence intensity 17%).

Figure 5 shows the pitch angle and generator speed responses of the baseline PI controller and the Multi-DLC emulator under DLC1.3. The RBF pitch command closely follows the baseline mean trend, although it continues to exhibit the small-amplitude, high-frequency oscillations already noted in the DLC1.1 test. These oscillations remain on the order of ±0.1–0.2° and do not shift the mean pitch appreciably: the Multi-DLC mean pitch angle (19.65°) deviates from the baseline (19.57°) by less than 0.5% (as shown in Table 4). Under the higher turbulence of DLC1.3, the generator speed fluctuates more widely around the rated value, as expected. The amplitude and pattern of these fluctuations are comparable between the two controllers, confirming that the RBF emulator does not introduce additional speed excursions.

The descriptive statistics in Table 4 further indicate that the mean and range of all three variables remain close to the baseline, with deviations within 1% for the means and no spurious extreme values.

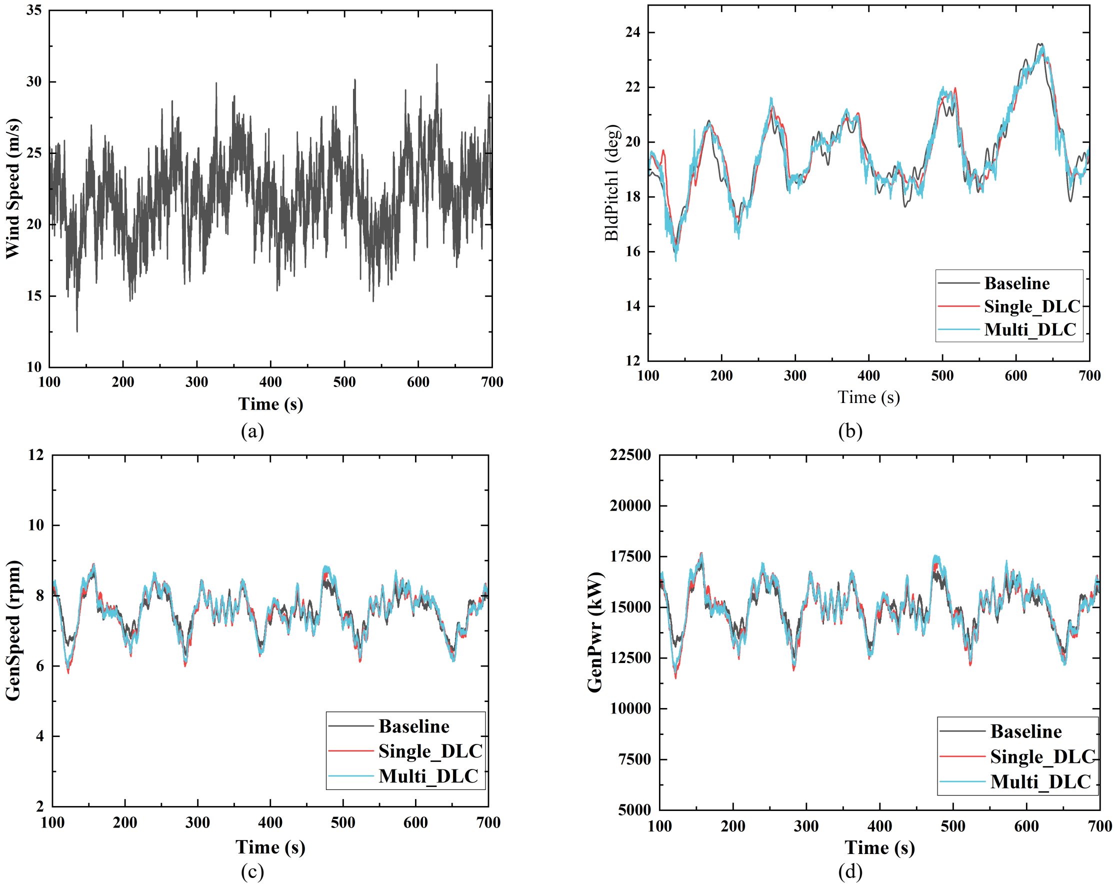

Figure 4. Responses comparison of the baseline OpenFAST PI controller (dark grey), the Single-DLC RBF emulator (red), and the Multi-DLC RBF emulator (blue) over a 600-second turbulent wind segment (DLC1.1, mean 22 m/s, TI = 2.5%). (a) Hub-height wind speed. (b) Collective pitch angle. (c) Generator speed. (d) Generator power.

Table 3. Statistical summary of pitch, generator speed, and power responses over 600 s of DLC1.1 turbulent wind (mean 22 m/s, TI = 2.5%).

Variable	Controller	Mean	Std Dev	Min	Max
Pitch angle (°)	Baseline	19.605	1.458	15.980	23.600
	Single-DLC	19.683	1.447	16.203	23.237
	Multi-DLC	19.639	1.501	15.647	23.517
Generator speed (rpm)	Baseline	7.658	0.468	6.305	8.745
	Single-DLC	7.536	0.609	5.790	8.915
	Multi-DLC	7.554	0.610	5.937	8.901
Generator power (kW)	Baseline	15016	928	12510	17350
	Single-DLC	14952	1208	11488	17688
	Multi-DLC	14988	1210	11779	17660

In summary, the closed-loop tests under both DLC1.1 and DLC1.3 demonstrate that the RBF emulator faithfully reproduces the OpenFAST CPC behavior across different wind regimes. The Single-DLC configuration provides accurate emulation under normal turbulence. The Multi-DLC configuration extends this fidelity to extreme turbulence conditions without architectural changes. The high-frequency pitch oscillations observed in the Multi-DLC model persist under both low- and high-turbulence conditions; however, they are effectively filtered by the rotor inertia and do not propagate into the generator speed or power responses. These results confirm that the proposed surrogate is suitable for fast simulation and control prototyping.

Table 4. Statistical summary of pitch, generator speed, and power responses over 600 s of DLC1.3 extreme gust conditions (mean 22 m/s, TI = 17%).

Variable	Controller	Mean	Std Dev	Min	Max
Pitch angle (°)	Baseline	19.57	2.33	14.08	24.50
	Multi-DLC	19.65	2.24	14.28	25.11
Generator speed (rpm)	Baseline	7.54	0.83	5.00	9.27
	Multi-DLC	7.52	0.98	4.97	9.96
Generator power (kW)	Baseline	14955.37	1684.10	8888.00	18400.00
	Multi-DLC	14921.46	1947.20	9856.07	19767.01

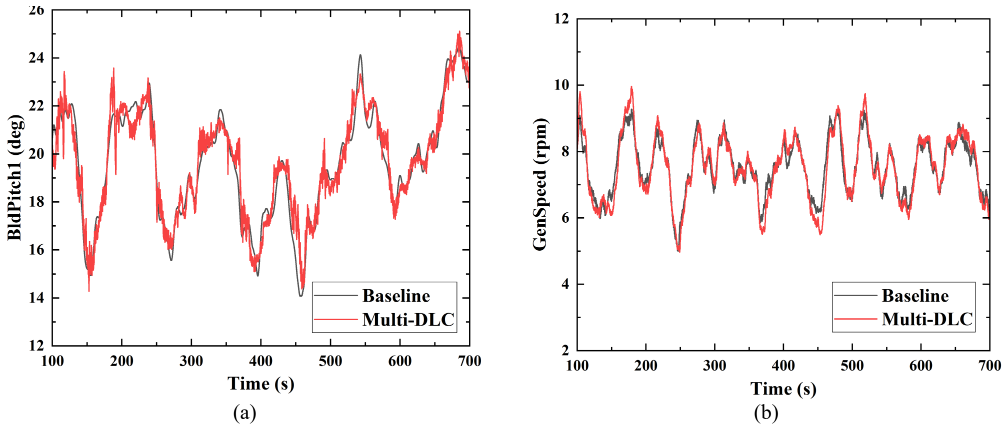

Figure 5. Responses comparison under DLC1.3 (mean 22 m/s, TI = 17%) for the Multi-DLC configuration. (a) Collective pitch angle. (b) Generator speed.

4. Conclusions and Future Works

This paper presents a data-driven emulation framework for collective pitch control (CPC) of wind turbines using radial basis function (RBF) networks and sliding time-window feature extraction. The model was trained with OpenFAST simulations under DLC1.1, DLC1.3, and DLC1.6 conditions, considering two configurations: Single-DLC (only DLC1.1) and Multi-DLC (all three load cases).

The results showed an accurate reproduction of the original controller, achieving offline metrics of $R^2 > 0.99$, RMSE < 0.5°, and Pearson correlation > 0.996, with negligible overfitting. Dynamic validation under DLC1.1 (22 m/s and TI = 2.5%) demonstrated that the emulator closely tracked the baseline PI controller. The Single-DLC configuration produced a smoother response, while the Multi-DLC setup introduced minor high-frequency oscillations that did not affect generator speed or power output. Although these oscillations are negligible for electromechanical performance, their potential long-term impact on pitch actuator fatigue and maintenance requires further quantification via power spectral density analysis, which is left for future work. Under the more severe DLC1.3 condition (TI = 17%), the Multi-DLC emulator maintained high fidelity and robustness to extreme turbulence.

Overall, the results indicate that a lightweight RBF network trained even on a single load case can provide a viable substitute for the original CPC, while incorporating multiple DLCs improves generalization under harsher operating conditions.

Future work will focus on extending the approach to individual pitch control (IPC) for asymmetric load mitigation, incorporating online or incremental learning to adapt to turbine degradation and changing dynamics. Additional research directions include exploring more advanced neural architectures.

ACKNOWLEDGMENT

This work has been partially supported by the Spanish Ministry of Science and Innovation under the projects MCI/AEI/FEDER number PID2024-155653OB-C21.

References

Allen, C., Viscelli, A., Dagher, H., Goupee, A., Gaertner, E., Abbas, N., Hall, M., & Barter, G. (2020). Definition of the UMaine VolturnUS-S Reference Platform Developed for the IEA Wind 15-Megawatt Offshore Reference Wind Turbine (NREL/TP--5000-76773). National Renewable Energy Laboratory (NREL), Golden, CO (United States); Univ. of Maine, Orono, ME (United States). https://doi.org/10.2172/1660012

Bai, D., Wang, B., Tan, J., & Wang, W. (2025). Research on Pitch Control for Floating Wind Turbines. 2025 44th Chinese Control Conference (CCC), 629–633. https://doi.org/10.23919/CCC64809.2025.11179773

Espinoza, V., Ormaza, C., Vidal, Y., & Tutivén, C. (2025). Optimizing Collective Pitch Control in Wind Turbines Using Deep Reinforcement Learning and OpenFAST Simulation⁎. IFAC-PapersOnLine, 1st IFAC Workshop on Smart Energy System for Efficient and Sustainable Smart Grids and Smart Cities - SENSYS 2025, 59(9), 151–156. https://doi.org/10.1016/j.ifacol.2025.08.128

Gambier, A. (2021). Pitch Control of Three Bladed Large Wind Energy Converters—A Review. Energies, (23), 8083. https://doi.org/10.3390/en14238083

IEC 61400-1:2019. (n.d.). Retrieved September 18, 2025, from https://webstore.iec.ch/en/publication/26423

Jeong, M., Loth, E., Qin, C., Phadnis, M., Pusch, M., & Pao, L. (2025). Operational Pitch Actuation Dynamics for Offshore Wind Turbines Ranging from 5 to 50 MW. Wind Energy, 28(2), e2975. https://doi.org/10.1002/we.2975

Kane, M. B. (2020). Machine Learning Control for Floating Offshore Wind Turbine Individual Blade Pitch Control. 2020 American Control Conference (ACC), 237–241. https://doi.org/10.23919/ACC45564.2020.9147912

Muñoz-Palomeque, E., Sierra-García, J. E., & Santos Peñas, M. (2024). Técnicas de control inteligente para el seguimiento del punto de máxima potencia en turbinas eólicas. https://doi.org/10.4995/riai.2024.21097

Routray, A., Sivakumar, N., Hur, S., & Bang, D. (2023). A Comparative Study of Optimal Individual Pitch Control Methods. Sustainability, 15(14), 10933. https://doi.org/10.3390/su151410933

Russell, A. J., Collu, M., McDonald, A. S., Thies, P. R., Keane, A., & Quayle, A. R. (2024). LIDAR-assisted feedforward individual pitch control of a 15 MW floating offshore wind turbine. Wind Energy, 27(4), 341–362. https://doi.org/10.1002/we.2891

Serrano, C., Sierra-Garcia, J.-E., & Santos, M. (2022). Hybrid Optimized Fuzzy Pitch Controller of a Floating Wind Turbine with Fatigue Analysis. Journal of Marine Science and Engineering, 10(11), 1769. https://doi.org/10.3390/jmse10111769

Sierra-García, J. E., & Santos, M. (2021). Lookup Table and Neural Network Hybrid Strategy for Wind Turbine Pitch Control. Sustainability, 13(6), 3235. https://doi.org/10.3390/su13063235

Velino, J., Kang, S., & Kane, M. B. (2022). Machine Learning Control for Floating Offshore Wind Turbine Individual Blade Pitch Control. Journal of Computing in Civil Engineering, 36(6), 04022034. https://doi.org/10.1061/(ASCE)CP.1943-5487.0001043

Zheng, S., Wu, L., Wang, L., Wang, L., & Hong, Y. (2025). Improving Operational Stability of Large-Scale Wind Turbines Through Data-Driven Model Predictive Individual Pitch Control. IEEE Transactions on Sustainable Energy, 16(4), 2445–2457. https://doi.org/10.1109/TSTE.2025.3554029