energies

MDPI

Article

The Improved Model Predictive Pitch Control Method for Wind Turbines Based on LiDAR

Zhihao Jin and Dongfei Fu *

The College of Engineering, Ocean University of China, Qingdao 266100, China; 21230911006@stu.ouc.edu.cn
* Correspondence: fudongfei@ouc.edu.cn

Abstract

This paper presents a LiDAR-informed adaptive-cost nonlinear model predictive control (NMPC) strategy for wind turbine pitch regulation. The proposed method uses a reinforcement learning (RL) agent as a supervisory cost-shaping module that adjusts the weights in the NMPC cost function. The pitch command is obtained from the constrained NMPC optimizer, which preserves the physical prediction model, actuator limits, and receding-horizon solution structure. LiDAR-derived preview wind-speed information is used as an estimate of the incoming disturbance and is introduced into both the prediction model and the agent state. This design helps the controller account for wind-field variation over the prediction horizon and adapt the relative emphasis on power regulation, load mitigation, and pitch-action smoothness. Compared with feedforward PID (FF-PID) and fixed-weight feedforward NMPC (FF-NMPC) controllers, the proposed controller shows stronger adaptability under abrupt and stochastic wind variations in OpenFAST-MATLAB/Simulink co-simulations.

Keywords: wind turbine; LiDAR; feedforward control; model predictive control; reinforcement learning

1. Introduction

With the global energy mix undergoing a low-carbon transition and the rapid growth of large-scale renewable energy deployment, wind power has become one of the most important forms of green energy [1]. Advances in wind energy conversion technologies have driven the vigorous development of the wind power industry. The power-generation efficiency of wind turbines has continuously improved, and turbine design and manufacturing technologies have gradually matured [2]. The emergence of large-scale wind turbines [3,4] has enabled wind turbines to operate in low-wind-speed regions and improved their adaptability to diverse terrains and climatic conditions. During wind turbine operation, the wind speed within the rotor-swept area is often non-uniform due to factors such as wind shear, tower shadow, and yaw misalignment. Such non-uniform inflow can impose unbalanced structural loads on the blades, thereby shortening the service life of the turbine [5–7].

Conventional pitch angle control methods typically employ PID controllers, in which power regulation and speed control are achieved through empirical parameter tuning [8,9]. However, classical linear controllers often fail to deliver satisfactory dynamic performance under turbulent inflow, gust disturbances, and extreme operating conditions [10–12]. These controllers have limited capability to address nonlinear system dynamics and cannot

Check for updates

Academic Editors: Rui Xiong and Francesco Castellani

Received: 28 January 2026
Revised: 22 April 2026
Accepted: 28 April 2026
Published: 1 May 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

systematically account for trade-offs between system states and control objectives; consequently, power fluctuations and mechanical loads may remain significant. To enhance wind turbine control performance, optimal control theory has been introduced into the pitch control domain [13,14]. Kwansu [15] incorporated LQR control as a feedforward component into a PI control structure to optimize overall performance. Haoyuan Wen [16] developed an LQR-enhanced disturbance-adaptive controller. Engels et al. [17] proposed an LQR-based variational approach that combines an extended Kalman filter with gain scheduling based on aerodynamic sensitivity and state-feedback gains to approximate nonlinear control, improving control performance under complex dynamic conditions. Nevertheless, LQR remains limited in handling constraints and nonlinear system dynamics.

Extensive research has also been conducted on MPC-based pitch systems [18,19]. El-Baklish et al. [20] developed an aeroelastic wind turbine rotor model by extending Hamilton's principle and incorporating the coupled flapwise and torsional blade motions, and subsequently designed a nonlinear model predictive pitch controller. Ahmed et al. [21] proposed an explicit model predictive output-feedback controller that can effectively handle system constraints while reducing online computational burden; moreover, the controller exhibits strong robustness against unmodeled uncertainties.

Feedforward control represents another important approach for improving pitch control performance [22,23]. In pitch regulation, future wind-speed disturbances are often the primary drivers of load variations and power oscillations. A nacelle-mounted forward-looking LiDAR [24], as a high-precision non-contact wind-measurement sensor, can acquire incoming wind-speed field in advance, thereby providing predictive disturbance inputs to the controller. Yamaguchi et al. [25] combined radar feedforward control with nacelle acceleration feedback, effectively reducing the fluctuating loads on the first-order mode of the tower. LiDAR-based feedforward control methods [26,27] introduce preview wind information to support model predictive control in adjusting control actions proactively, substantially improving the response to gust disturbances, reducing structural load fluctuations, and enhancing the robustness of power output. Existing studies indicate that control strategies incorporating wind-speed preview outperform conventional feedback control and provide a more forward-looking solution for wind turbine pitch control [28].

Table 1 positions the proposed controller against representative NMPC, FF-NMPC, and RL-based pitch control strategies in terms of LiDAR preview, explicit constraint handling, and online parameter tuning.

Table 1. Comparison of representative pitch control strategies.

Controller Type	LiDAR Preview	Explicit Constraints	Online Parameter Tuning	Relative Limitation or Advantage
Fixed-weight NMPC	No	Yes	No	Lacks preview wind information and online objective adaptation.
Fixed-weight FF-NMPC	Yes	Yes	No	Uses preview wind estimates, while the balance among control objectives follows offline weights.
RL-based pitch control	No	No	Yes	Provides adaptive parameter adjustment, while hard constraints are not handled in an intuitive and explicit form.
Proposed adaptive-cost RL-NMPC	Yes	Yes	Yes	Combines preview wind estimates, online parameter tuning, and constrained NMPC to coordinate power stability, load mitigation, and pitch-action smoothness.

"Yes" indicates that the corresponding feature or capability is present for the controller type, while "No" indicates that it is absent.

This study develops a LiDAR-informed adaptive-cost RL-NMPC strategy for wind turbine pitch control. The main contributions are summarized as follows:

1. A LiDAR-preview-informed NMPC prediction model is constructed. The rotor-equivalent preview wind-speed sequence is embedded in the prediction horizon to provide time-varying estimates of wind-speed-dependent aerodynamic variables, including the tip-speed ratio, power coefficient, and thrust coefficient. This formulation gives the prediction model access to incoming wind trends over the horizon and reduces reliance on a single current or mean wind-speed value.
2. A reinforcement-learning-based cost-shaping mechanism is introduced. The RL agent outputs the weighting parameters of the NMPC cost function, and the pitch command is computed by the NMPC optimizer under system and actuator constraints.
3. A LiDAR-informed fast-slow adaptive-cost framework is developed. The agent state combines turbine operating variables with preview wind-field statistics, the fast layer performs NMPC rolling optimization, and the slow layer updates the cost weights through PPO with weight-rate, KL-divergence, and dynamic-interval mechanisms. The method is validated in OpenFAST v3.5.5 [29] and MATLAB/Simulink2023b under EOG [30] and 17 m/s random wind conditions.

The limitations of this study are that the simulation validation is based on the NREL 5-MW reference wind turbine model and that no physical rig or field test has been conducted. Future work will incorporate real operational data from actual wind turbines for engineering refinement of the algorithm.

2. Mathematical Model

The wind turbine generator comprises three core modules: mechanical, electrical and control. The mechanical module includes the wind turbine blades, transmission mechanism and supporting structure. The electrical module integrates the generator and grid-connection equipment. The control module covers pitch adjustment, torque optimization and yaw positioning systems. Through the coordinated operation of these multiple modules, the system achieves an efficient closed-loop process of wind energy capture, mechanical transmission, and electrical conversion, ensuring stable output and load balance of the unit under dynamic wind conditions.

2.1. Wind Turbine System Model

The aerodynamic subsystem of a wind turbine rotor converts the kinetic energy of the incoming flow into mechanical energy. The aerodynamic power extracted by the rotor depends on wind speed, the power coefficient, and the rotor-swept area, where V denotes the effective wind speed, C_p is the power coefficient, and S represents the rotor-swept area. For a specific turbine design, a quasi-stationary C_p characteristic can be used in engineering calculations. Specifically, once the tip-speed ratio λ and the blade pitch angle β are specified, the corresponding C_p value can be directly obtained from a pre-established $C_p(\lambda, \beta)$ map. This approach is computationally efficient and is commonly referred to as the "lookup-table method", which is underpinned by extensive turbine-characteristic measurements and accumulated engineering experience. The C_p characteristic of the NREL 5-MW reference wind turbine is shown in Figure 1. The rotor aerodynamic torque T_r is given by

$$T_r = \frac{P_r}{\omega_r} = \frac{1}{2}\rho\pi R^2 V^3 \frac{C_p}{\lambda} \tag{1}$$

where ρ is the air density, R is the rotor radius.

Aerodynamic thrust acting on the wind turbine's rotor is given by

$$F_T = \frac{1}{2}\rho\pi R^2 V^2 C_t \tag{2}$$

where F_T denotes the rotor thrust and C_t is the thrust coefficient, expressed as a nonlinear function of the tip-speed ratio and the blade pitch angle, as illustrated in Figure 2.

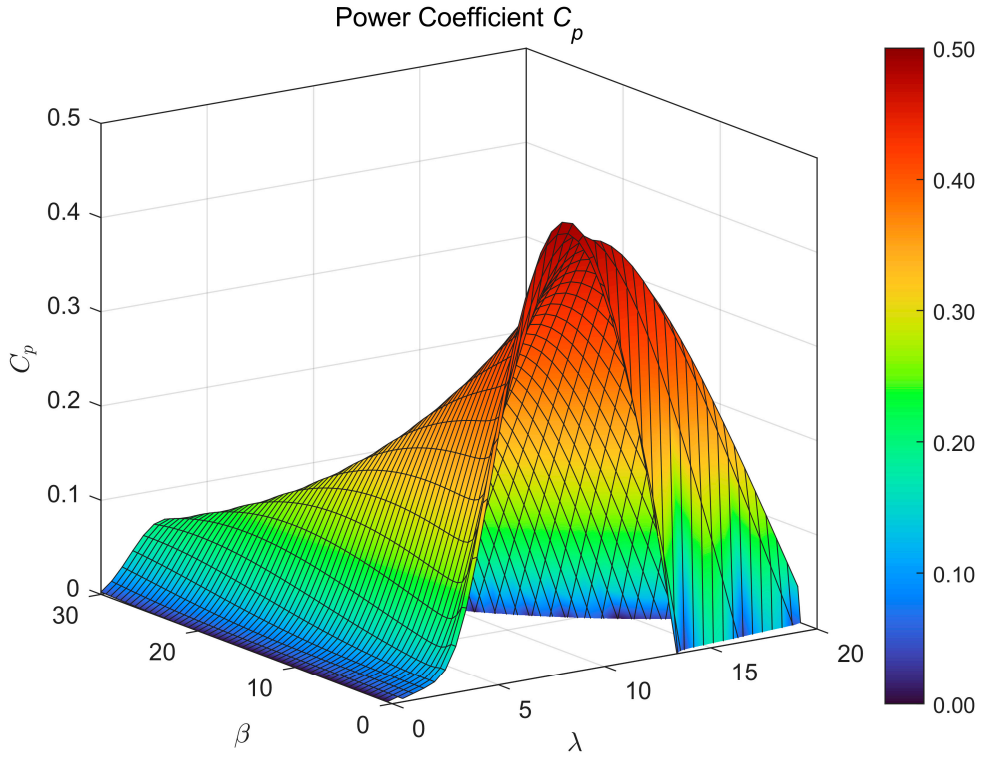


Figure 1. Schematic diagram of the power coefficient–pitch angle–tip-speed ratio relationship for a NREL 5-MW wind turbine.

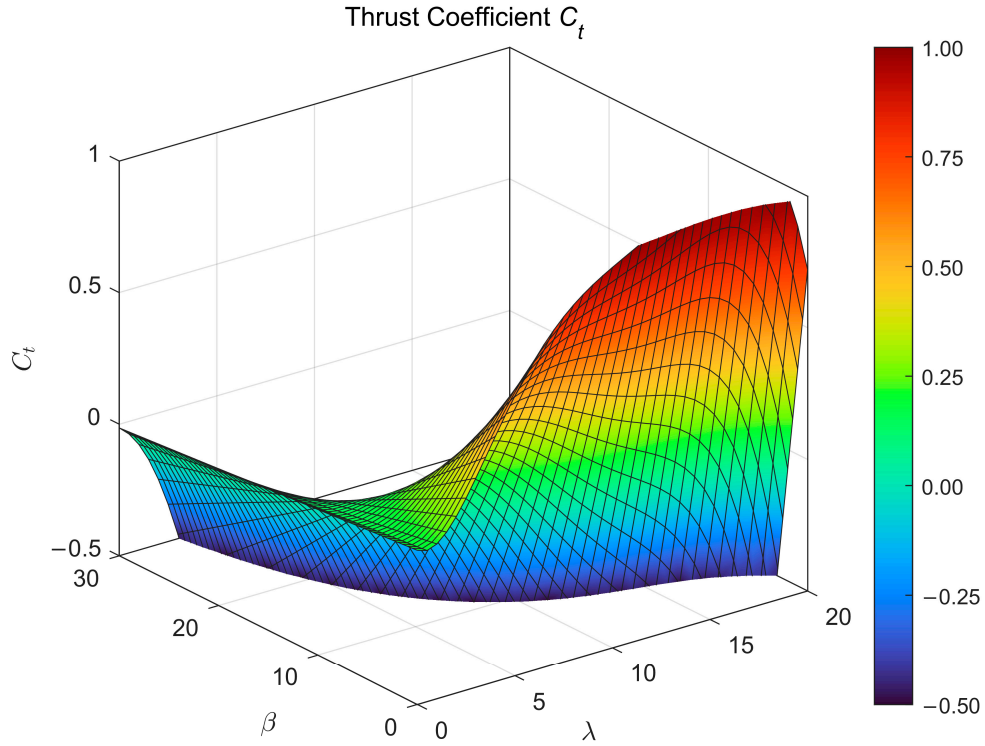


Figure 2. Schematic diagram of the thrust coefficient–pitch angle–tip-speed ratio relationship for a NREL 5-MW wind turbine.

Under wind loading, the rotor and tower of a wind turbine exhibit complex multi-order modal responses. In general, the flexible blades show relatively high modal damping, whereas the tower and drivetrain typically exhibit lower but non-negligible damping. In this study, attention is focused on the first fore–aft bending mode of the tower, since it typically dominates the structural response. By modeling the wind turbine as a mechanical system, the equation of motion can be used to describe the first-mode fore–aft tower displacement as follows:

$$m\ddot{x}_T = F_T - D_T\dot{x}_T - K_T x_T \tag{3}$$

where m is the total rotor (wind wheel) mass, D_T is the equivalent (modal) damping coefficient associated with the first fore–aft tower mode, K_T is the corresponding (modal) stiffness coefficient, and F_T denotes the aerodynamic thrust acting on the rotor.

2.2. Operating Principle of LiDAR

Based on Doppler frequency-shift detection, a nacelle-mounted Doppler LiDAR can measure the three-dimensional vector characteristics of the incoming wind field in real time. The system is designed in strict accordance with the IEC 61400-50-3:2022 standard issued by the International Electrotechnical Commission (IEC) [31], enabling accurate reconstruction of the 3D wind vector field within a range of 50–400 m upstream of the rotor hub. As shown in Figure 3, the laser transmitter adopts a dual-plane, four-beam spatial scanning architecture. The four beams have a horizontal divergence angle of 15° and a vertical divergence angle of 7.5°, and are labeled in a clockwise order relative to the LiDAR. The four-channel line-of-sight wind-speed measurements are fused, processed using a Kalman filter, and synchronously fed into the wind turbine main control system. This provides critical feedforward information for control optimization and improved energy capture efficiency, thereby enabling adaptive control based on wind-field perception.

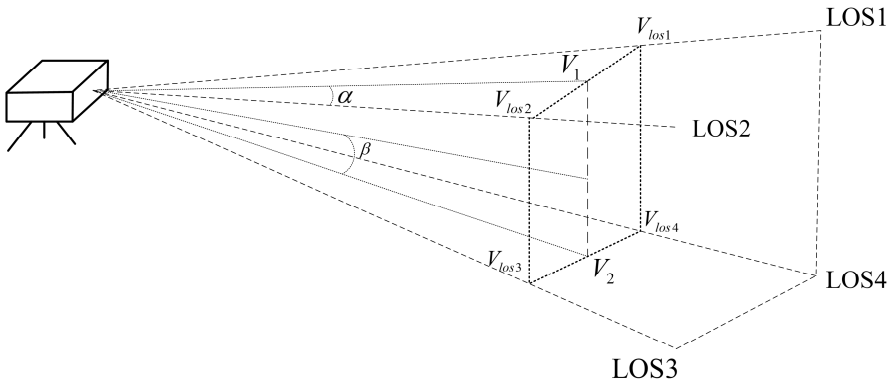


Figure 3. Schematic diagram of the four-beam method. Dashed lines indicate line-of-sight (LOS).

The atmospheric wind field above the wind turbine is assumed to satisfy a local homogeneity condition within a limited region, i.e., the flow is statistically steady and the turbulence properties are spatially uniform (homogeneous turbulence). This assumption is justified by the fact that the LiDAR measurement domain is sufficiently small compared with the characteristic scales of large-scale atmospheric motions; therefore, variations in the velocity gradient and turbulence intensity within this region can be neglected, which provides a simplified basis for wind-speed modeling.

Within the LiDAR measurement region, the measured wind speed at hub height can be regarded as the area-averaged wind speed over an equivalent (virtual) rotor disk, since the hub height coincides with the center height of the rotor plane and their wind-speed characteristics are therefore highly consistent. To further improve the accuracy of the

equivalent wind-speed estimate, the concept of the "upper–lower-beam composite wind speed" is introduced and defined as follows:

1. Upper-beam composite wind speed V_1: the resultant wind speed synthesized from the two laser beams in the upper scanning plane;
2. Lower-beam composite wind speed V_2: the resultant wind speed synthesized from the two laser beams in the lower scanning plane;
3. Rotor-equivalent wind speed V_{avg}: defined as the arithmetic mean of the upper- and lower-beam composite wind speeds, i.e.,

The wind speeds of the upper- and lower-beam pairs can be decomposed into a longitudinal component u and a lateral component v. The upper-plane composite wind speed V_1 is calculated as follows, where u_1 and v_1 denote the longitudinal and lateral components of the upper-beam pair, respectively:

$$\begin{cases} v_1 = \dfrac{V_{los2} - V_{los1}}{2 \sin \alpha \cos \beta} \\ u_1 = \dfrac{V_{los2} + V_{los1}}{2 \cos \alpha \cos \beta} \\ V_1 = \sqrt{v_1^2 + u_1^2} \end{cases} \tag{4}$$

where V_{los1} and V_{los2} are the wind speeds measured along Line-of-Sight 1 and 2. α is the horizontal angle between the two laser beams. β is the elevation angle of the beams relative to the horizontal plane.

Similarly, the lower-plane composite wind speed V_2 can be obtained from (5), where u_2 and v_2 are the longitudinal and lateral components of the lower-beam pair, respectively:

$$\begin{cases} v_2 = \dfrac{V_{los3} - V_{los4}}{2 \sin \alpha \cos \beta} \\ u_2 = \dfrac{V_{los3} + V_{los4}}{2 \cos \alpha \cos \beta} \\ V_2 = \sqrt{v_2^2 + u_2^2} \end{cases} \tag{5}$$

where V_{los3} and V_{los4} are the wind speeds measured along Line-of-Sight 3 and 4.

Let the distance between the wind-measurement center and the turbine rotor be d, and the mean wind speed be v_{avg}. Under the Taylor frozen-turbulence hypothesis, the LiDAR can provide preview information of the incoming wind acting on the rotor with a lead time of

$$T_{pre} = \frac{d}{v_{avg}} \tag{6}$$

The processed wind-speed information must further experience an additional delay before it becomes available to the control system, as illustrated in Figure 4. Accordingly, the effective preview time is given by

$$T = \frac{d}{v_{avg}} - 0.5T_{scan} - T_{filter} - \tau \tag{7}$$

where T_{filter} denotes the filtering-induced delay and T_{scan} is the LiDAR scanning delay. τ is the time constant of the pitch actuator.

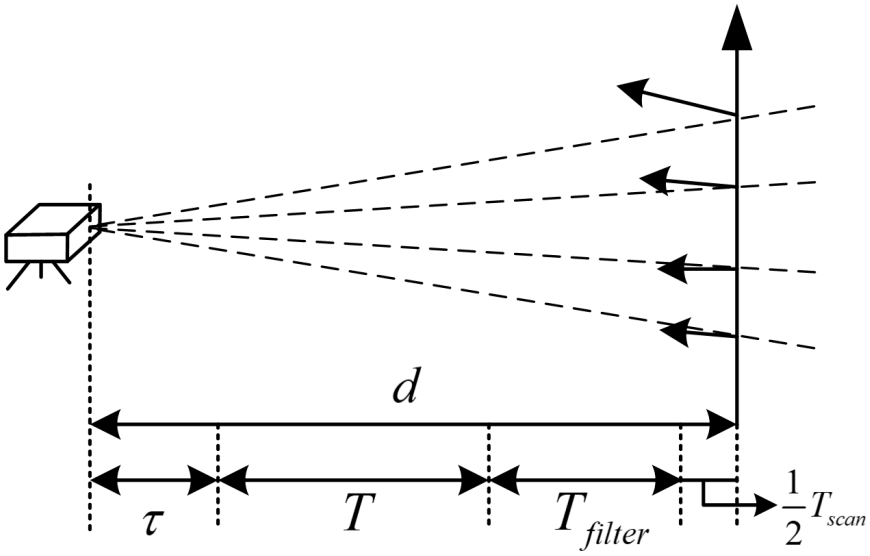


Figure 4. Schematic diagram of the rotor-equivalent wind speed. Dashed lines denote the LiDAR line-of-sight (LOS) beams, and solid arrows indicate the wind velocity vectors at different positions across the rotor plane.

3. Methodology

This section introduces the proposed method in a progressive order. First, the operating region relevant to pitch regulation is identified. Then, the LiDAR-preview-based NMPC prediction model and cost function are formulated. Finally, the PPO-based cost-weight adaptation and its coupling with the constrained NMPC optimizer are described in detail.

According to the operational characteristics of variable-speed wind energy conversion systems under different wind-speed conditions, power control is typically divided into four regions (see Figure 5). In Region 1, the wind speed is below the cut-in speed, and the turbine remains in the start-up and grid-connection preparation stage. Region 2 corresponds to the maximum energy capture region, where the turbine operates in variable-speed mode to achieve maximum power point tracking. In Region 3, the turbine enters the constant-power operating region, and the output power is regulated and stabilized by adjusting the blade pitch and related control actions. In Region 4, the wind speed exceeds the cut-out speed, and the turbine switches to a feathering protection mode and ceases power generation. Therefore, the primary operating region of pitch control is Region 3.

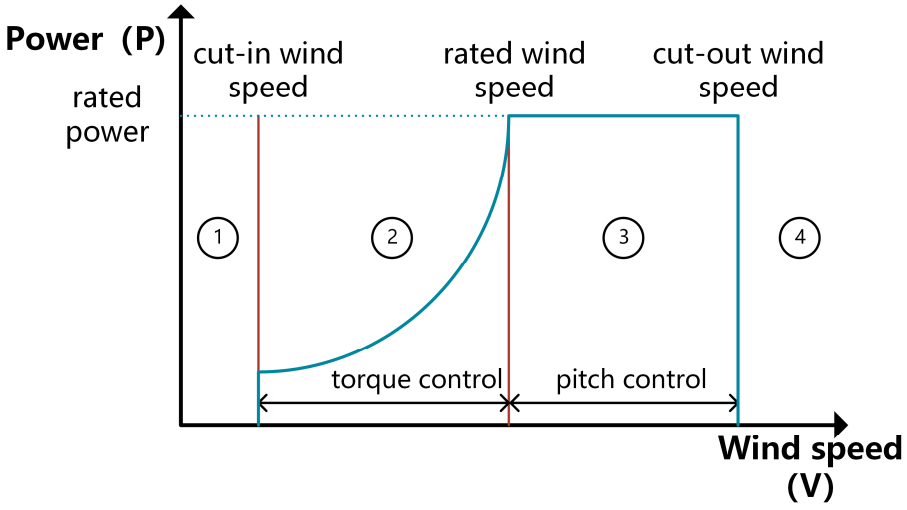


Figure 5. Schematic diagram of wind turbine operating regions.

3.1. Prediction Model of the NMPC Controller

Let the state vector of the wind turbine be defined as

$$x = \begin{bmatrix} \omega_r & \omega_g & \psi & x_T & \dot{x}_T \end{bmatrix}^{\mathrm{T}} \tag{8}$$

With the control input $u = \beta$, the continuous-time state-update relationship of the system is given by

$$\dot{x} = \begin{bmatrix} \frac{1}{J_r}[T_r(\omega_r, \beta, V) - (D_z + K_s)\psi] \\ \frac{(D_s + K_s)\psi}{N_g J_g} - \frac{T_g}{J_g} \\ \frac{1}{J_r}[T_r - (D_z + K_s)\psi] - \frac{1}{N_g}\left(\frac{(D_z + K_s)\psi}{N_g J_g} - \frac{T_g}{J_g}\right) \\ \dot{x}_T \\ \frac{1}{m}\left[F_T(\omega_r, \beta, V) - D_T\dot{x}_T - K_T x_T\right] \end{bmatrix} \tag{9}$$

The model is discretized using the fourth-order Runge–Kutta method, yielding

$$x_{k+1} = x_k + \frac{T_s}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{10}$$

where k_1, k_2, k_3, k_4 are the four intermediate slope vectors.

Within the model predictive control framework, the prediction model should reflect the main wind-speed variations that affect turbine dynamics. Here, the LiDAR preview sequence is used as a previewed exogenous wind sequence over the prediction horizon. The LiDAR preview sequence is treated as an estimated rotor-equivalent wind-speed trajectory affected by sensing and propagation uncertainty. It is used to update wind-speed-dependent aerodynamic terms in the NMPC prediction model, and the related aerodynamic quantities are interpreted as preview-informed estimates.

Let the current time instant be k, and the prediction horizon length be N_p. The LiDAR system can provide a predicted sequence of rotor-equivalent wind speed over the next N_p steps:

$$V_{pred}(k) = \left[V(k|k), \ldots, V\left(k + N_p - 1|k\right)\right]^{\mathrm{T}} \tag{11}$$

The tip-speed ratio λ is defined as the ratio of the blade-tip linear velocity to the incoming wind speed, and it varies in real time with wind speed and rotor speed. At the i-th step within the prediction horizon, a preview-informed estimate of the tip-speed ratio is given by

$$\lambda(k + i|k) = \frac{\omega_r(k + i|k) \cdot R}{V(k + i|k)} \tag{12}$$

where $\omega_r(k + i|k)$ denotes the rotor speed predicted at time k for the future instant $k + i$, and R is the rotor radius. Equation (12) indicates that λ is no longer a parameter determined solely by the current state; instead, it becomes a time-varying quantity that is tightly coupled with the predicted future wind speed.

After introducing the preview wind-speed sequence, the power coefficient C_p and the thrust coefficient C_t over the prediction horizon can be evaluated as preview-informed time-varying quantities:

$$\begin{cases} C_p(k + i|k) = f_{C_p}(\lambda(k + i|k), \beta(k + i|k)) \\ C_t(k + i|k) = f_{C_t}(\lambda(k + i|k), \beta(k + i|k)) \end{cases} \tag{13}$$

To meet the requirement of a learnable cost function, a differentiable model predictive control (diff-MPC) framework can be adopted, which imposes strict mathematical

requirements on the continuity and differentiability of the system model. A key advantage of diff-MPC is that the underlying optimization problem can be solved efficiently via gradient backpropagation; its convergence behavior and computational efficiency depend directly on the accuracy and smoothness of the Jacobian and Hessian matrices of the system dynamics and the cost function. Therefore, all aerodynamic parameters involved in the model must admit continuous and differentiable analytical expressions, so that the optimization algorithm can compute gradient directions stably and accurately throughout the iterative process. To this end, we employ empirical formulations for the NREL 5-MW wind turbine as follows:

$$\begin{cases} C_p(\lambda,\beta) = c_1\left(\frac{c_2}{\lambda} - c_3\beta - c_4\right)e^{-\frac{c_5}{\lambda}} + c_6\lambda \\ C_t(\lambda,\beta) = -\sum_{i=1}^{5} d_i \cdot \lambda^i + d_{10} - \beta\left(d_6 + \lambda^2 d_7 - \lambda^3 d_8 + \lambda^4 d_9\right) \end{cases} \tag{14}$$

where c_1~c_6 and d_1~d_{10} are constants.

Since $\lambda(k+i|k)$ varies with the predicted wind speed, C_p and C_t also change dynamically over the prediction horizon. The aerodynamic torque T_r and the aerodynamic thrust F_T depend directly on the squared or cubic terms of the wind speed. Accordingly, Equations (1) and (2) can be reformulated at the i-th step of the prediction horizon as

$$\begin{cases} T_r(k+i|k) = \frac{1}{2}\rho\pi R^2[V(k+i|k)]^3 \frac{C_p(\lambda(k+i|k),\beta(k+i|k))}{\omega_r(k+i|k)} \\ F_T(k+i|k) = \frac{1}{2}\rho\pi R^2[V(k+i|k)]^2 C_t(\lambda(k+i|k),\beta(k+i|k)) \end{cases} \tag{15}$$

These preview-informed formulations include the estimated wind-speed variation over the prediction horizon, giving NMPC access to incoming wind trends for proactive pitch optimization. Based on these preview-informed formulations, the nonlinear prediction model of the wind turbine can be expressed in a unified form as follows:

$$\begin{aligned} &x(k+i+1|k) = f_d(x(k+i|k), u(k+i|k), V(k+i|k)) \\ &i = 0, 1, \ldots, N_p - 1 \end{aligned} \tag{16}$$

Here, the definitions of x and u remain unchanged. In the model function $f_d(\cdot)$, all wind-speed-dependent terms are evaluated with the preview sequence $V(k+i|k)$ when available, yielding a wind-preview-informed time-varying prediction model over the horizon.

3.2. Cost Function Design

In NMPC for wind turbines, the design of the cost function is central to optimizing control performance. To achieve stable and efficient operation, the cost function should account for multiple objectives, including tracking of the output variables, tracking of the control inputs, and smoothness of the control actions. Specifically, the rotor-speed tracking cost is formulated as follows:

$$J_{\omega,k} = \sum_{i=1}^{N_p} (\omega_{r,k+i\,|\,k} - \omega_{\text{ref}})^2 \tag{17}$$

where ω_{ref} is the rated rotor speed.

Tower-top displacement velocity cost:

$$J_{T,k} = \sum_{i=1}^{N_p} \dot{x}_{T,k+i|k}^2 \tag{18}$$

Control-rate (input increment) cost:

$$J_{\Delta\beta,k} = \sum_{i=0}^{N_c-1} (\Delta\beta_{k+i\,|\,k})^2 \tag{19}$$

By jointly accounting for the above terms, the overall cost function can be written as

$$J_k = \lambda_{1\theta,k} J_{\omega,k} + \lambda_{2\theta,k} J_{T,k} + \lambda_{3\theta,k} J_{\Delta\beta,k} \tag{20}$$

where $\lambda_{1\theta,k}$, $\lambda_{2\theta,k}$ and $\lambda_{3\theta,k}$ are weighting coefficients that quantify the relative importance of the individual performance terms. Here, N_p is the prediction horizon, N_c is the control horizon, $\omega_{r,k+i|k}$ is the predicted rotor speed at step $k+i$ based on information available at time k. ω_{ref} is the reference rotor speed. $\dot{x}_{T,k+i|k}$ is the predicted tower-top displacement velocity, $\Delta\beta_{k+i|k}$ is the pitch angle increment, and k denotes the current sampling instant.

Similar to conventional model predictive control, the control problem can be formulated as the following optimization:

$$\begin{aligned} &\left(X_k^*, U_k^*\right) = \text{argmin}_{X_k, U_k} J_k(\lambda_{k\theta}) \\ &\text{s.t. } x_{k|k} = x_k,\ x_{k+i+1|k} = f_d(x_{k+i|k}, u_{k+i|k}),\ i = 0, \ldots, N_p - 1, \\ &\quad x_{k+i|k} \in \mathcal{X},\ i = 1, \ldots, N_p,\ u_{k+i|k} \in \mathcal{U},\ i = 0, \ldots, N_p - 1 \end{aligned} \tag{21}$$

In the adaptive-cost setting, Equation (21) is treated as an implicit NMPC module parameterized by the cost-weight vector. Let ϕ denote the PPO actor parameters and let $w_k = g_\theta(s_k) = [\lambda_{1\theta,k}, \lambda_{2\theta,k}, \lambda_{3\theta,k}]^{\mathrm{T}}$ be the weight vector generated from the normalized agent state. s_k denotes the state input of the agent. For the given turbine state x_k, LiDAR preview sequence $V_{pred,k}$, and weight vector w_k, the constrained NMPC problem in Equation (21) returns the optimal predicted state-control trajectory Z_k^*. According to the receding-horizon principle, the applied pitch command is the first element of the optimal control sequence and can be written as the implicit control law:

$$u_k = \beta_k = q_{\text{NMPC}}\left(x_k, V_{pred,k}, w_k\right) \tag{22}$$

Here, q_{NMPC} denotes the control mapping induced by the constrained NMPC optimizer. Therefore, the learning layer changes only the objective weights of the NMPC problem, while the physical prediction model, actuator limits, and feasible set in Equation (21) remain unchanged. For the weight-learning layer, let L_k be a one-step performance loss used to evaluate the closed-loop response. Since L_k depends on the actor parameters through the composite mapping $\theta \to w_k \to u_k \to x_{k+1}$, the chain rule gives

$$\frac{dL_k}{d\theta} = \frac{\partial L_k}{\partial u_k} \frac{\partial q_{\text{NMPC}}}{\partial w_k} \frac{\partial g_\theta(s_k)}{\partial \theta} \tag{23}$$

The first factor is obtained from the performance loss, the last factor is obtained by standard neural-network backpropagation, and the middle factor is the local sensitivity of the NMPC optimizer with respect to the cost weights. This local NMPC sensitivity can be obtained by differentiating the Karush–Kuhn–Tucker conditions of the local SQP/iLQR approximation and recursively propagating the resulting sensitivity terms along the prediction horizon.

3.3. Agent Design

In the proposed RL-NMPC structure, the PPO agent tunes the NMPC cost weights using turbine states and LiDAR-derived preview wind statistics. The state vector is six-dimensional and is constructed as follows:

$$s_k = \left[V_k, \omega_{r,k}, \beta_k, F_{base,k}, \mu_{v,fut}, \sigma_{v,fut}\right]^{\mathrm{T}} \tag{24}$$

where V_k is the effective wind speed at the current time step at the rotor hub height, $\omega_{r,k}$ is the rotor speed at the current time step, β_k is the pitch angle at the current time step,

and $F_{base,k}$ represents the tower-base force. The forecasted mean wind speed for the future prediction horizon is given by

$$\mu_{v,fut} = \frac{1}{N_p}\sum_{i=1}^{N_p} V_{k+i} \tag{25}$$

where $\mu_{v,fut}$ represents the arithmetic average of the predicted wind speeds, indicating the overall level of future wind energy. The forecasted wind speed variability is quantified by the standard deviation:

$$\sigma_{v,fut} = \sqrt{\frac{1}{N_p - 1}\sum_{i=1}^{N_p}\left(V_{k+i} - \mu_{v,fut}\right)^2} \tag{26}$$

where $\sigma_{v,fut}$ represents the sample standard deviation of the future wind speed forecast, indicating the variability of future wind speeds. Including these state variables enables the controller to capture both the level and variability of future wind speed, thereby improving power-tracking control.

The output of the agent in the NMPC cost function corresponds to the weighting parameters $\lambda_{1\theta}, \lambda_{2\theta}, \lambda_{3\theta}$, which are associated with the terms in Equation (20). To ensure the parameter values fall within an appropriate range and are convenient for optimization, the Sigmoid function is employed for the nonlinear mapping:

$$\lambda_{i\theta} = w_{\min} + (w_{\max} - w_{\min})\sigma(z_i) \tag{27}$$

where $w_{\min} = 0.1$ and $w_{\max} = 10$ represent the lower and upper bounds of the weighting parameters, respectively, and $\sigma(\cdot)$ is the standard Sigmoid activation function, which ensures that the weights stay within the feasible range. Because the mapping remains differentiable, it supports efficient backpropagation during end-to-end training. The agent operates as a self-tuning component that continuously adjusts the NMPC parameters based on the system's evolving state and environment. To reduce the influence of inconsistent physical dimensions, all elements of s_k are normalized by the mean and standard deviation obtained from the training rollouts.

The design of the reward function is closely related to the control objectives above the rated wind speed range, namely power stability and load reduction. The reward function is designed as follows:

$$R_k = -\exp\left(\left(\frac{P_t - P_{ref}}{200}\right)^2\right) - \exp\left(\left(\frac{F_T - \overline{F}_T}{200}\right)^2\right) - 10\sqrt{(\Delta\beta)^2} \tag{28}$$

where P_t denotes the generator output power at time instant t, P_{ref} is the power reference value, and its value is obtained from the rated power. F_T represents the tower-base force, and $\overline{F}$ denotes a mean value that is updated at each time step.

During PPO training, rollouts are collected from the closed-loop RL-NMPC simulation. At each environment step, the agent observes s_k, samples $a_k \sim \pi_{\boldsymbol{\theta}}(\cdot|s_k)$, maps it to $\lambda_{i\theta}$ using Equation (27), and passes $\lambda_{i\theta}$ to the NMPC optimizer. The optimizer computes the feasible pitch command, and the OpenFAST–Simulink environment returns the reward and the next state. The transition

$$z_k = (s_k,\ a_k,\ \lambda_{i\theta},\ r_k,\ s_{k+1},\ \log \pi_{\boldsymbol{\theta}}(a_k|s_k)) \tag{29}$$

is stored in the rollout buffer. After T_m samples are collected, the actor and critic are updated using the standard PPO clipped objective with generalized advantage estimation. Here, T_m is the adaptive rollout length before the m-th PPO update. You can find the RL hyperparameters and training settings in Table 2.

Table 2. RL hyperparameters and training settings [32].

Parameter	Value
Number of Layers In the Actor Network	2
Number of Neurons In the Hidden Layers of the Actor	512
Number of Layers In the Critic Network	2
Number of Neurons in the Hidden Layers of the Critic	512
Rollout Length	1024
Batch Size	128
Gradient Clipping	1.0
Discount Factor	0.98
GAE Lambda Coefficient	0.9
Clipping Parameter	0.12
Learning Rate (Actor/Critic)	$2 \times 10^{-4}/1 \times 10^{-4}$
Random Seed	7
Entropy Coefficient	0.04
Value Function Loss Coefficient	0.5

The PPO settings listed in Table 2 define the baseline training configuration. The KL-divergence threshold used for early stopping is 0.02.

3.4. LiDAR-Informed Adaptive-Cost RL-NMPC Framework

The proposed LiDAR-informed adaptive-cost RL-NMPC framework combines preview wind information, PPO-based cost-weight scheduling, and constrained NMPC optimization. The LiDAR module provides the rotor-equivalent preview wind-speed sequence and its statistics. The PPO agent maps the normalized turbine state and preview features to the NMPC cost weights. The NMPC layer then computes the pitch command by solving the constrained finite-horizon optimization problem.

The fast control layer is based on nonlinear model predictive control. At each sampling instant, it solves the constrained NMPC problem in Equation (21) and applies the first element of the optimal pitch sequence to the turbine:

$$\beta_k = \beta^*_{k|k} \tag{30}$$

The slow learning layer is based on PPO and updates the mapping from the agent state to the NMPC cost weights. The agent therefore works as an upper-level objective allocator connected to the turbine state, the LiDAR preview statistics, and the NMPC parameters.

To regulate the online learning process, the rollout length before PPO updating is adjusted adaptively. Let T_m denote the number of closed-loop environment steps collected before the m-th PPO update. After the m-th PPO update is accepted, the updated policy is applied over an evaluation window, and a recent-performance index is computed from tracking and control-related errors in this window. Let $\mathrm{P}(\tau)$ denote this post-update evaluation index, which is used to assess the system's control performance over the last τ steps, is computed as a combination of errors such as tracking and control discrepancies. The performance error is defined as

$$e_P(\tau) = \frac{1}{\tau} \sum_{i=k-\tau+1}^{k} \left(\frac{P_{e,i} - P_{er}}{P_{er}} \right)^2 \tag{31}$$

The loading fluctuation indicator is defined as

$$e_M(\tau) = \frac{1}{\tau} \sum_{i=k-\tau+1}^{k} \left(\frac{M_{t,i} - \overline{M}_t}{\sigma_M} \right)^2 \tag{32}$$

where M_t denotes the tower base bending moment, and $\overline{M}_t$ is the mean of M_t.

Therefore, the overall cost function $\mathrm{P}(\tau)$ is defined as

$$P(\tau) = \alpha_1 e_P(\tau) + \alpha_2 e_M(\tau) \tag{33}$$

where $\alpha_1 = 3000$ and $\alpha_2 = 1$ are weighting coefficients.

A smaller value of $\mathrm{P}(\tau)$ indicates better recent closed-loop performance under the updated policy. When the post-update performance is satisfactory, T_m is increased for the next update cycle to reduce unnecessary policy updates and computational overhead. When the performance index deteriorates, T_m is shortened to allow more frequent policy improvement:

$$T_{m+1} = \begin{cases} \min\left(T_{\max},\ \kappa_{\uparrow} T_m\right) & \mathrm{P}(\tau) \le \rho_{low} \\ \max\left(T_{\min},\ \kappa_{\downarrow} T_m\right) & \mathrm{P}(\tau) \ge \rho_{\mathrm{high}} \\ T_m, & \rho_{\mathrm{low}} < \mathrm{P}(\tau) < \rho_{\mathrm{high}} \end{cases} \tag{34}$$

where $\rho_{low} = 1.8$ and $\rho_{high} = 2.5$ represent the boundaries for good and deteriorating performance, respectively. $\kappa_{\uparrow}$ and $\kappa_{\downarrow}$ are adjustment factors.

To smooth the applied NMPC cost-weight sequence while preserving bounded optimization preferences, a rate constraint is introduced for the accepted weight updates:

$$\left\| w_{m+1} - w_m \right\|_{\infty} \le \gamma \left\| w_m \right\|_{\infty} \tag{35}$$

where $\gamma = 0.05$ is the rate limit factor. If the new output weight w_{new} violates the rate constraint, scaling is performed:

$$w_{m+1} = w_m + \min\left(1, \frac{\gamma \left\| w_m \right\|_{\infty\infty}}{\left\| w_{new} - w_m \right\|}\right)(w_{new} - w_m) \tag{36}$$

Equation (36) projects the candidate update direction onto the infinity-norm ball centered at w_m with radius $\gamma \left\| w_m \right\|_{\infty}$. This limiter smooths excessively abrupt accepted weight changes, while the current actor mapping can still adjust the NMPC weights according to the measured state within the permitted bound.

To limit the magnitude of PPO policy updates, the KL-divergence constraint is written as the empirical average KL-divergence over the rollout batch $\mathcal{B}_m$:

$$\overline{D}_{\mathrm{KL}}^{(m)} = \frac{1}{|\mathcal{B}_m|} \sum_{s \in \mathcal{B}_m} D_{\mathrm{KL}}\left(\pi_{\theta_{\mathrm{old}}}(\cdot|s) \parallel \pi_{\theta}(\cdot|s)\right) \le \delta_{\max} \tag{37}$$

where $\mathcal{B}_m$ is the rollout batch used for the m-th policy update, and $\delta_{\max} = 0.02$ is the maximum allowable KL-divergence change.

The adaptive update interval, weight-rate constraint, and empirical KL-divergence bound regulate update frequency, accepted weight variation, and policy shift, respectively, thereby supporting smooth interaction between learning and NMPC.

In this formulation, online adaptation is treated as a bounded scheduling change in the NMPC objective. The prediction model, actuator limits, and feasible input set are unchanged; only the cost weights in the objective function are adjusted. Let $J_N(x; w)$ denote the finite-horizon optimal value of the NMPC problem at state x under the cost-weight vector w. Within the validated operating region, if $J_N(x; w)$ is locally Lipschitz continuous with respect to w, then for the same current state x_k and two consecutive accepted weight vectors w_k and w_{k-1}, there exists a local constant $L_w > 0$ such that

$$|J_N(x_k; w_k) - J_N(x_k; w_{k-1})| \le L_w |w_k - w_{k-1}| \tag{38}$$

Equation (38) means that the change in the finite-horizon objective caused by an accepted weight update is bounded by the size of that weight update. Therefore, the rate constraint in Equations (35) and (36) helps avoid abrupt reshaping of the NMPC objective while keeping the prediction model, actuator constraints, and receding-horizon optimization structure unchanged. This local regularity statement is considered within the validated operating region, under feasible NMPC solutions, positive bounded cost weights, bounded actor outputs, and the PPO training diagnostics in Section 3.3 indicating stable

moving-average episode return and bounded mean KL-divergence. The LiDAR-informed adaptive-cost RL-NMPC procedure proposed in this paper is summarized in Algorithm 1.

Algorithm 1. LiDAR-informed adaptive-cost RL-NMPC procedure

Inputs: x_k, s_k, $V_{pred,k}$, w_{k-1}, parameter sets Ω_{PPO}, Ω_{NMPC}
Output: β_k

1. Acquire the turbine state x_k and the LiDAR preview sequence $V_{pred,k}$.
2. Construct the normalized agent state s_k using turbine measurements and preview wind statistics.
3. Evaluate the actor mapping. During training, sample $a_k \sim \pi_{\boldsymbol{\theta}}(\cdot|s_k)$. During deterministic validation or deployment, use the policy mean. Map a_k to the candidate cost-weight vector.
4. Apply the weight bounds and rate constraint to obtain the accepted NMPC weight w_k.
5. Solve the constrained NMPC problem with x_k, $V_{pred,k}$, w_k, and apply the first pitch command.
6. During training, store the transition in the rollout buffer. When T_m samples have been collected and the learning condition is satisfied, update PPO using the training procedure in Section 3.3.
7. Apply the updated policy over an evaluation window, compute $\mathrm{P}(\tau)$, and update the next rollout length T_{m+1}.
8. Set $k \leftarrow k+1$ and repeat from Step 1.

4. Simulation Results

To verify the controller's performance under different wind speed conditions, this section conducts joint simulation using the professional simulation software MATLAB/Simulink. The controller is integrated with the NREL 5-MW wind turbine model provided by OpenFAST. The pitch actuator speed is limited to -8~$+8^\circ$/s, and the pitch angle is constrained between 0° and 30°. The remaining simulation model parameters are shown in Table 3.

Table 3. Wind turbine system parameters.

Parameter	Value
Rated Power (MW)	5.25
Blade Radius (m)	63
Number of Blades	3
Cut-in Wind Speed (m/s)	4
Cut-out Wind Speed (m/s)	25
Rated Wind Speed (m/s)	11.4
Rated Rotor Speed (rpm)	12.1

For the NMPC controller employed in this work, the sampling time T_s is set to 0.1 s, the control horizon is set to $N_c = 2$, with a prediction horizon of $N_p = 10$. In this paper, the iLQR (Iterative Linear Quadratic Regulator) framework is used to decompose the optimization problem into a series of quadratic programming (QP) subproblems. When pitch angle/pitch rate hard constraints are present, the PNQP (Projected Newton Quadratic Programming) solver is used to solve the QP subproblem with upper and lower bounds. This solving strategy balances the nonlinear modeling accuracy of NMPC with real-time solving capability, adapting to the fast control requirements of the wind turbine under varying wind speed conditions. To verify the accuracy and effectiveness of the dynamic

cost-weighting model predictive control method proposed in this study, comparative simulations are conducted against FF-PID, FF-NMPC, and RLMPC. The LiDAR processes the incoming wind speed data and transmits it to the pre-compensated controller. To quantitatively evaluate the performance of different control algorithms, this paper selects the Root Mean Square Error (RMSE) of electric power output and Damage Equivalent Load (DEL) based on tower-base fore-aft bending moment as the core evaluation indicators. The RMSE of electric power output is used to characterize the deviation between the actual electric power output of the wind turbine and the rated electric power. A smaller value indicates better stability of electric power output and better power-tracking performance of the control algorithm:

$$\mathrm{RMSE}(P) = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(P_t - P_{rated})^2} \tag{39}$$

DEL converts the random alternating tower-base fore-aft bending moment into an equivalent constant amplitude alternating load, which is used to evaluate the fatigue damage degree of the tower structure. A smaller value indicates a smaller fatigue load on the tower and a better load suppression effect of the control algorithm:

$$\mathrm{DEL} = \left(\frac{\sum_{i=1}^{n} n_i M_i^m}{N_{ref}} \right)^{\frac{1}{m}} \tag{40}$$

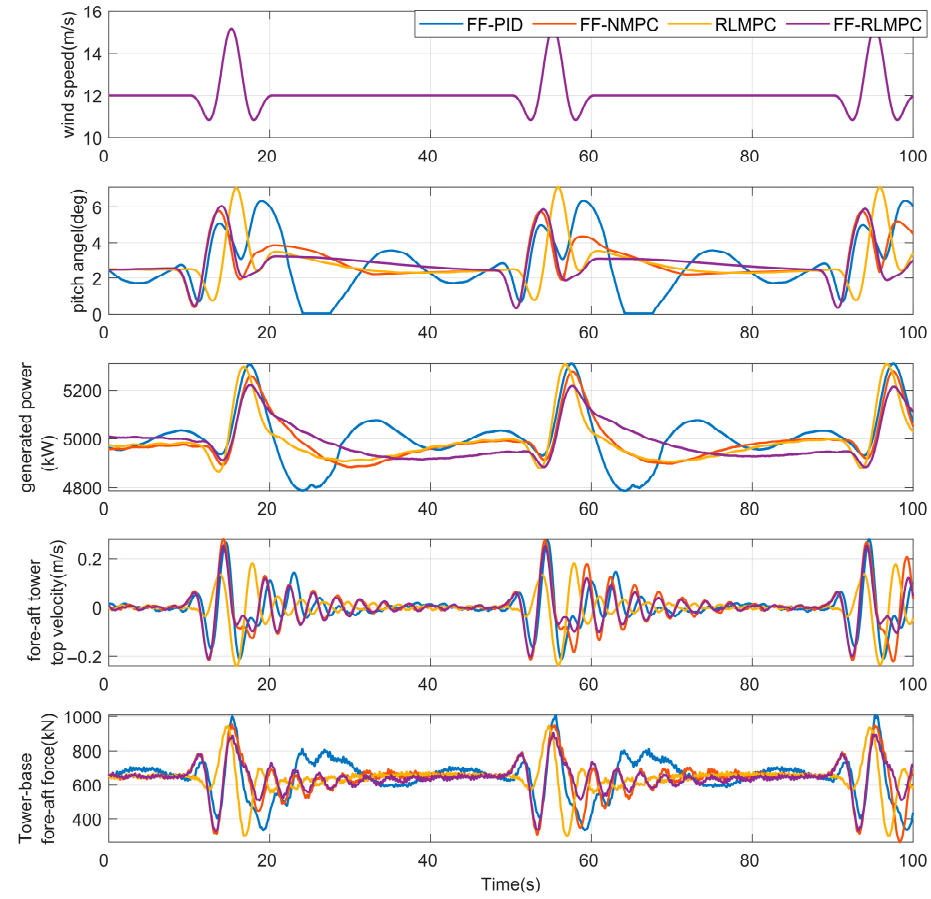
where n is the number of load levels with different amplitudes in the load spectrum. n_i is the number of cycles of the i-th amplitude load. M_i represents the fore-aft bending moment load of the tower base with the i-th amplitude. N_{ref} is the reference number of cycles, taking 10^7 cycles. $m = 3$ is the fatigue exponent of the S-N curve of the tower structure for the wind turbine steel tower. To clarify the correspondence between OpenFAST output channels, physical quantities, and their respective applications in the proposed control framework, Table 4 summarizes the specific mapping relationships. Specifically, different tower-related physical quantities derived from OpenFAST are assigned to distinct roles, as detailed in the following table.

Table 4. Definition and application of OpenFAST tower-related output channels.

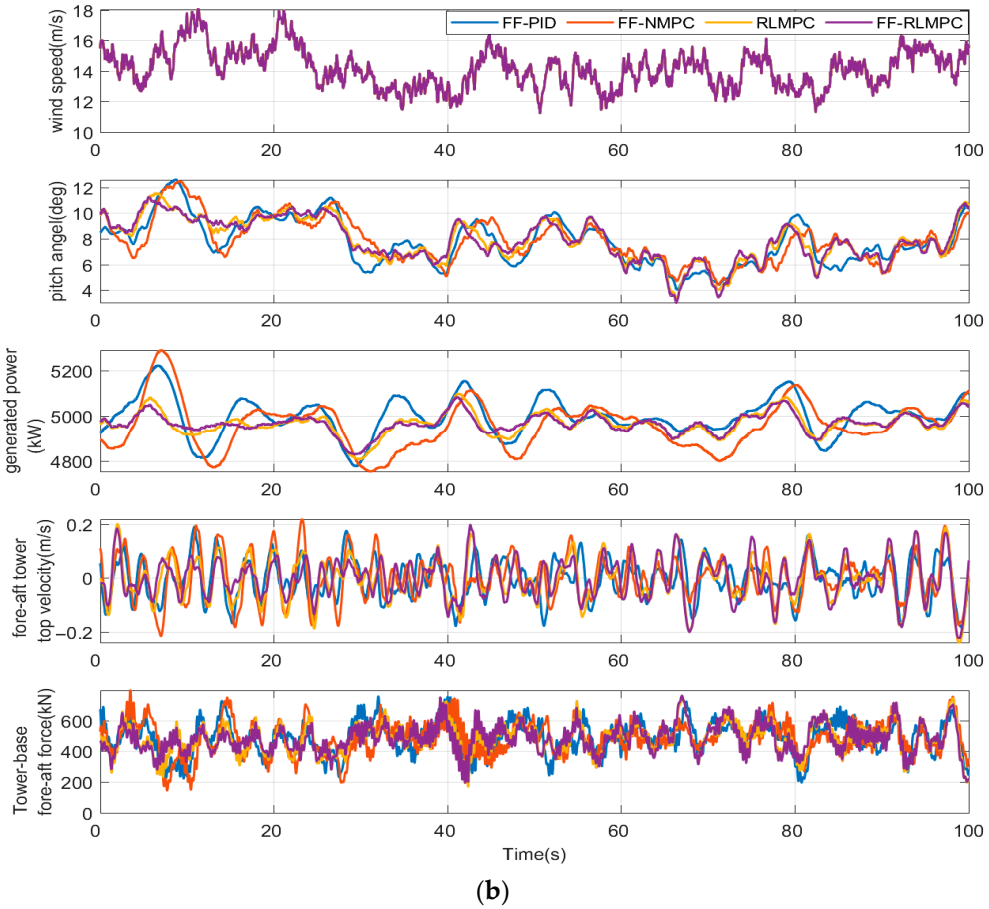
OpenFAST Output Channel	Physical Quantity	Application
TwrBsFxt	Tower-base fore-aft force	Controller state input
TwrBsMxt	Tower-base fore-aft bending moment	Fatigue damage evaluation

The PID parameters are tuned using the empirical trial-and-error method with repeated adjustments to balance speed tracking, dynamic response, and system stability. The weighting parameters of the NMPC controller are determined through extensive simulation comparisons and debugging. Based on the normalization of cost function variables, a reasonable balance is achieved among speed tracking, load mitigation, and actuator smoothness, with corresponding weights of 2, 0.3, and 0.05. Controllers using feedforward preview are supplied with the same LiDAR preview; RLMPC uses current rotor-equivalent wind-speed input by design.

The validation uses one EOG case and three turbulent wind cases. The EOG case is used to evaluate transient response under abrupt wind-speed variation. Turbulent case 1, turbulent case 2, and turbulent case 3 are used to evaluate controller performance under stochastic turbulent inflow. The four wind-speed waveforms are shown in Figure 6.

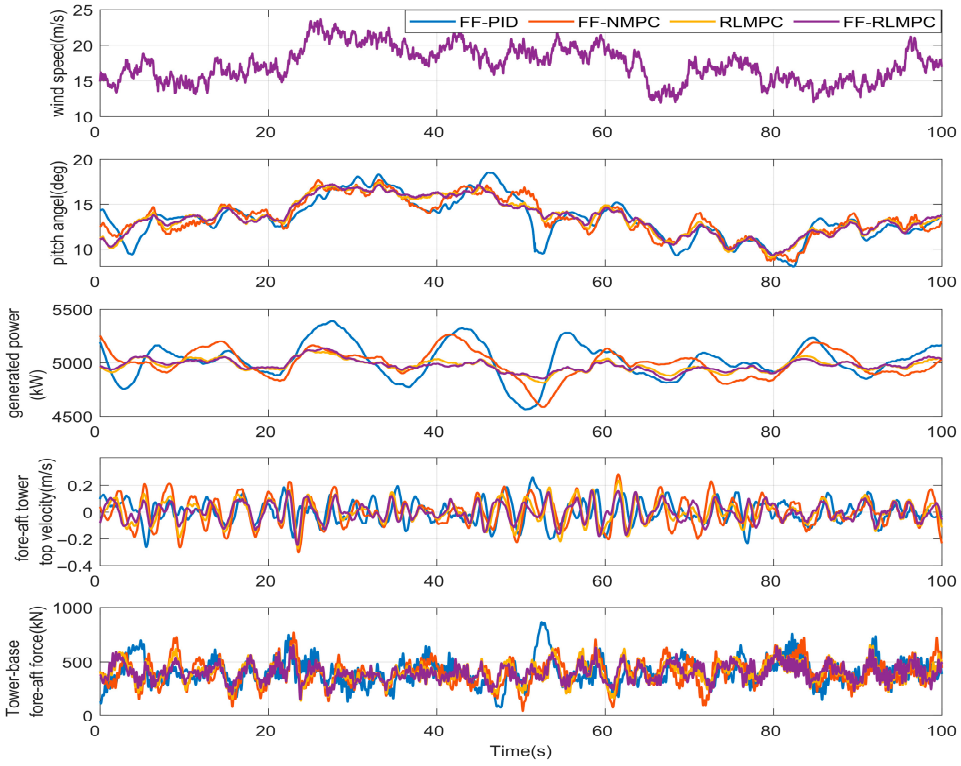


(**a**)

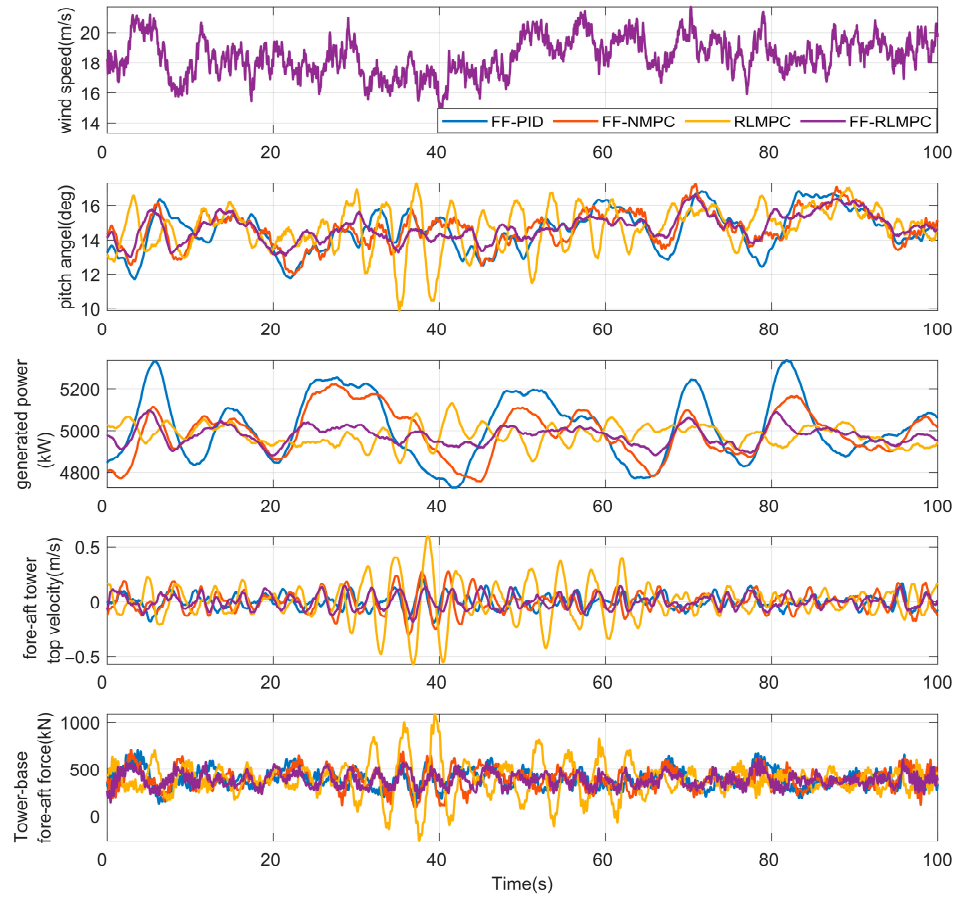


(**b**)

Figure 6. *Cont*.



(**c**)



(**d**)

Figure 6. Wind-speed waveforms of the four validation cases. (**a**) EOG. (**b**) Turbulent case 1. (**c**) Turbulent case 2. (**d**) Turbulent case 3.

Four controllers are compared: FF-PID, FF-NMPC, RLMPC, and FF-RLMPC. FF-PID is used as the engineering feedforward baseline. FF-NMPC uses LiDAR preview and fixed NMPC weights. RLMPC uses adaptive-cost weights and current rotor-equivalent wind-speed input. FF-RLMPC uses LiDAR preview, adaptive-cost weights, and constrained NMPC optimization. This comparison structure is used to evaluate the effects of LiDAR preview, adaptive-cost-weight scheduling, and constrained predictive optimization.

4.1. Hyperparameter Sensitivity Under 17 m/s Turbulent Wind

To evaluate whether the proposed RL-NMPC controller is overly sensitive to PPO training hyperparameters, a compact sensitivity analysis was conducted under a 17 m/s turbulent wind condition. Since the purpose of this test is to examine the robustness of the learning configuration rather than to compare different controller structures, all cases use the same LiDAR-informed RL-NMPC framework, turbine model, reward function, NMPC settings, and turbulent wind condition. Only selected PPO hyperparameters are varied.

Considering the computational cost of RL-NMPC training, four representative settings are considered, as listed in Table 5. H0 is the baseline setting used in the main simulations. H1 reduces the actor/critic learning rates to examine the influence of a conservative update step. H2 increases the actor/critic learning rates to test whether overly aggressive parameter updates degrade training stability. H3 moderately increases the PPO clipping parameter from 0.12 to 0.16 to examine the influence of a less conservative policy-update range. All remaining hyperparameters are kept identical to the baseline setting.

Table 5. Hyperparameter settings for sensitivity analysis under 17 m/s turbulent wind.

Case	Purpose	Actor/Critic Learning Rate	PPO Clip
H0	Baseline	$2 \times 10^{-4}/1 \times 10^{-4}$	0.12
H1	Conservative learning rate	$1 \times 10^{-4}/5 \times 10^{-5}$	0.12
H2	Aggressive learning rate	$4 \times 10^{-4}/2 \times 10^{-4}$	0.12
H3	Moderately larger policy-update range	$2 \times 10^{-4}/1 \times 10^{-4}$	0.16

The reward convergence curves are shown in Figure 7. For clarity, the raw episode reward is shown only for the baseline case, while the 15-episode moving-average curves are plotted for all four cases. The gray curve denotes the raw average reward per step of H0, and the colored curves denote the corresponding smoothed reward trends.

Based on the recorded training data in Figure 7, the baseline case H0 reaches the highest final smoothed reward. The conservative learning-rate case H1 improves more slowly and remains below the baseline within the same training horizon, indicating that an excessively small learning rate may delay policy improvement. The aggressive learning-rate case H2 shows faster early improvement than H1, but its final reward remains lower than H0, suggesting that overly large actor–critic updates may disturb the cost-weight scheduling process in the RL-NMPC framework. The moderately larger clipping case H3 remains close to the baseline, with a slightly lower final reward and slightly stronger variation. This indicates that a moderate increase in the PPO clipping range does not fundamentally change the convergence trend, but it may slightly reduce update conservativeness.

The reward-based metrics are summarized in Table 6. The final reward is taken from the last value of the 15-episode moving-average curve. The convergence episode is defined here as the first episode after which the smoothed reward remains within 0.5 of its final value. The last-50-episode standard deviation is computed from the smoothed reward curve and is used only as a simple indicator of late-stage fluctuation.

Overall, the results suggest that the selected baseline setting provides a reasonable balance between convergence speed and update stability. H1 is more conservative but

slower, H2 is more aggressive but does not provide a better final reward, and H3 remains close to the baseline while introducing slightly stronger variation.

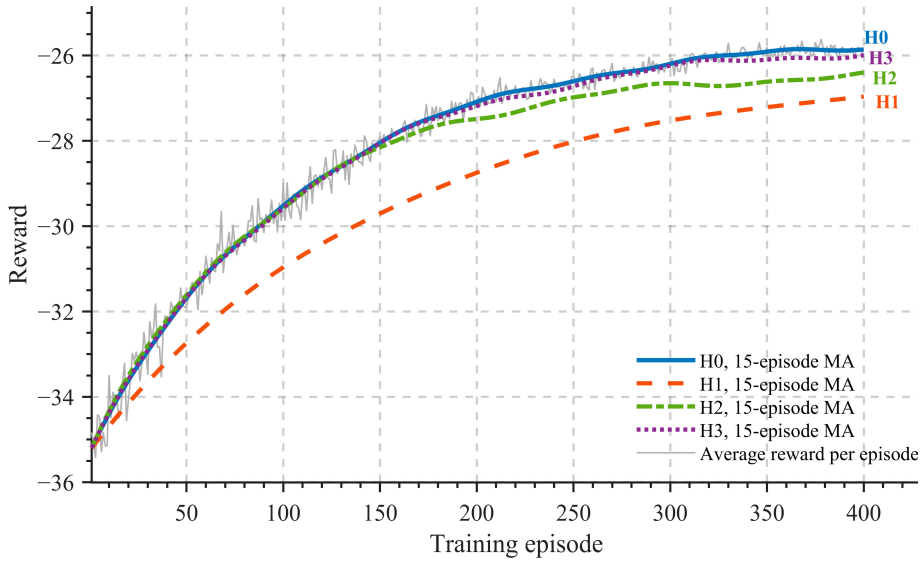


Figure 7. Reward convergence under different PPO hyperparameter settings.

Table 6. Reward-based sensitivity metrics.

Case	Final 15-Episode Avg. Reward	Reward Improvement	Convergence Episode	Last-50-Episode Std.
H0	−25.858	9.336	282	0.014
H1	−26.964	8.235	309	0.072
H2	−26.399	8.794	263	0.061
H3	−25.990	9.202	270	0.021

4.2. Controller Performance Under EOG and Turbulent Wind Cases

This section compares the four controllers under one EOG case and three turbulent wind cases. The quantitative results are listed in Table 7. For each wind case, the same wind-speed input is used for all controllers, which ensures a fair comparison among FF-PID, FF-NMPC, RLMPC, and FF-RLMPC.

Table 7. Quantitative comparison of four controllers under EOG and turbulent wind cases.

Wind Case	Controller	Peak-to-Trough Power Output	Power RMSE	Pitch-Angle Std	DEL
EOG	FF-PID	521	111.648	1.596	0.235701
EOG	FF-NMPC	392	90.112	1.205	0.195828
EOG	RLMPC	442	93.687	1.118	0.258352
EOG	FF-RLMPC	260	61.002	1.131	0.196096
Turbulent case 1	FF-PID	486	91.507	1.886	0.315587
Turbulent case 1	FF-NMPC	549	113.707	1.791	0.317792
Turbulent case 1	RLMPC	295	59.979	1.695	0.315961
Turbulent case 1	FF-RLMPC	253	52.528	1.676	0.313127
Turbulent case 2	FF-PID	622	155.368	1.187	0.467886
Turbulent case 2	FF-NMPC	465	112.145	1.019	0.373559
Turbulent case 2	RLMPC	288	52.660	1.245	0.417625
Turbulent case 2	FF-RLMPC	214	49.389	0.769	0.354151
Turbulent case 3	FF-PID	588	115.208	2.050	0.369300
Turbulent case 3	FF-NMPC	531	116.032	1.795	0.321150
Turbulent case 3	RLMPC	283	57.325	1.697	0.337333
Turbulent case 3	FF-RLMPC	244	51.358	1.675	0.297750

Peak-to-trough power output represents the absolute peak-to-trough difference in electrical power output, in kW; Mean Power Output RMSE represents the Root-Mean-Square Error of the rated power output, in kW.

As shown in Table 7, FF-RLMPC achieves the lowest peak-to-trough power output and the lowest power RMSE in all four wind cases. This indicates that the proposed controller provides the strongest power-fluctuation suppression among the four methods. In the EOG case, FF-RLMPC reduces the power RMSE by 45.36%, 32.30%, and 34.89% compared with FF-PID, FF-NMPC, and RLMPC, respectively. Its DEL is close to that of FF-NMPC, indicating that the transient power-regulation improvement is achieved while maintaining a comparable tower-base fatigue-load level.

For the three turbulent cases, FF-RLMPC also obtains the lowest DEL values. In turbulent case 1, FF-RLMPC reduces the power RMSE by 42.60%, 53.80%, and 12.42% compared with FF-PID, FF-NMPC, and RLMPC, respectively. In turbulent case 2, the corresponding reductions are 68.21%, 55.96%, and 6.21%. In turbulent case 3, the corresponding reductions are 55.42%, 55.74%, and 10.41%. These results show that FF-RLMPC maintains improved power regulation under different turbulent inflow sequences.

For tower-base DEL, FF-RLMPC reduces the value by 0.78%, 1.47%, and 0.90% in turbulent case 1 compared with FF-PID, FF-NMPC, and RLMPC, respectively. In turbulent case 2, the corresponding reductions are 24.31%, 5.20%, and 15.20%. In turbulent case 3, the corresponding reductions are 19.37%, 7.29%, and 11.73%. The DEL results indicate that FF-RLMPC also improves fatigue-load mitigation under turbulent inflow.

The reduction ratio of FF-RLMPC relative to each baseline controller is calculated as

$$R^{(b)} = \frac{M_b - M_{\mathrm{FF-RLMPC}}}{M_b} \times 100\% \tag{41}$$

where M_b denotes the metric value of the baseline controller under the same wind case. Since smaller values of power RMSE and DEL indicate better performance, a positive reduction ratio represents an improvement of FF-RLMPC. The reduction ratios are shown in Figure 8, where Figure 8a corresponds to power RMSE and Figure 8b corresponds to tower-base DEL.

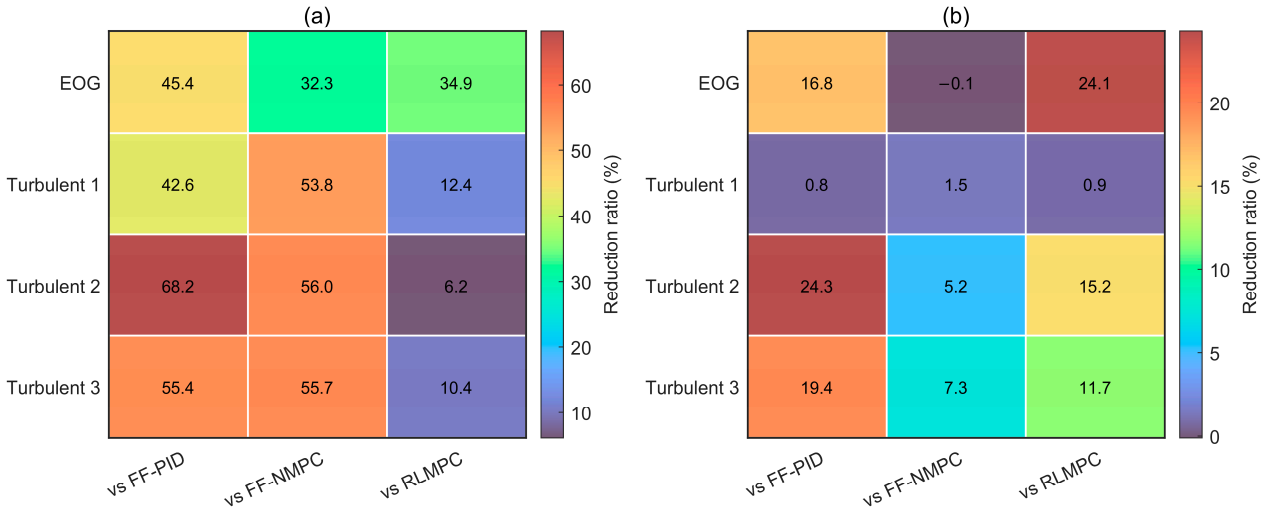


Figure 8. Reduction ratios of FF-RLMPC relative to the baseline controllers. (**a**) Power RMSE reduction ratio. (**b**) Tower-base DEL reduction ratio.

Overall, the comparison among the four controllers clarifies the role of each module. FF-NMPC reflects the effect of LiDAR-preview-based NMPC with fixed weights. RLMPC reflects the effect of adaptive-cost-weight scheduling with current wind-speed input. FF-RLMPC combines LiDAR preview, adaptive-cost-weight scheduling, and constrained NMPC optimization. The results show that FF-RLMPC improves both power-regulation accuracy and fatigue-load mitigation under EOG and turbulent inflow, while also maintaining smoother pitch behavior in the turbulent cases.

5. Conclusions

This study developed a LiDAR-informed adaptive-cost RL-NMPC strategy for wind turbine pitch control under above-rated wind conditions. The proposed framework integrates rotor-equivalent preview wind-speed information, PPO-based cost-weight adjustment, and constrained nonlinear model predictive control. In this structure, the reinforcement learning module updates the weighting coefficients of the NMPC cost function, while the prediction model, actuator constraints, and receding-horizon optimization process are retained.

The proposed FF-RLMPC controller was evaluated using the NREL 5-MW reference wind turbine in OpenFAST–MATLAB/Simulink under one EOG case and three turbulent wind cases. The simulation results show that FF-RLMPC achieves the lowest-power RMSE among the four compared controllers in all validation cases. In the EOG case, the power RMSE is reduced by 45.36%, 32.30%, and 34.89% compared with FF-PID, FF-NMPC, and RLMPC, respectively. For the three turbulent cases, the corresponding reductions are 42.60–68.21% relative to FF-PID, 53.80–55.96% relative to FF-NMPC, and 6.21–12.42% relative to RLMPC. These results indicate that the combination of LiDAR preview and adaptive-cost-weight adjustment is effective in reducing power-output fluctuations under the tested wind conditions.

The tower-base fatigue-load results also show improvements under turbulent inflow. In the three turbulent cases, FF-RLMPC obtains the lowest DEL values among the compared controllers. The DEL reductions are 0.78–1.47% in turbulent case 1, 5.20–24.31% in turbulent case 2, and 7.29–19.37% in turbulent case 3, depending on the baseline controller used for comparison. In the EOG case, the DEL value of FF-RLMPC remains close to that of FF-NMPC, suggesting that the transient power-regulation improvement is achieved without a clear increase in tower-base fatigue load.

Overall, the results demonstrate that the proposed FF-RLMPC controller can improve power regulation while providing load-mitigation benefits in the turbulent cases considered in this study. The largest performance gains are observed in power RMSE reduction, whereas the DEL improvement varies with the wind condition and comparison baseline. Since the present validation is based on simulation of the NREL 5-MW reference turbine, future work will extend the assessment to more diverse inflow conditions, LiDAR measurement uncertainty, and experimental or field-test data to further examine the engineering applicability of the proposed control strategy.

Author Contributions: Conceptualization, Z.J.; Methodology, Z.J.; Software, Z.J.; Validation, Z.J.; Formal analysis, Z.J.; Writing—original draft, Z.J.; Writing—review and editing, D.F.; Supervision, D.F.; Project administration, D.F.; Funding acquisition, D.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Key R&D Program of Shandong Province (Major Science and Technology Innovation Project), China (Grant No. 2023CXGC010408).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. IRENA. *World Energy Transitions Outlook 2022: 1.5 °C Pathway*; (IRENA) International Renewable Energy Agency: Abu Dhabi, United Arab Emirates, 2022.
2. Amano, R.S. Review of wind turbine research in 21st century. *J. Energy Resour. Technol.* **2017**, *139*, 050801. [CrossRef]
3. Dai, J.; Li, M.; Chen, H.; He, T.; Zhang, F. Progress and challenges on blade load research of large-scale wind turbines. *Renew. Energy* **2022**, *196*, 482–496. [CrossRef]

4. Antonysamy, R.P.; Joo, Y.H. Power maximization and regulation of the super-large wind turbine system using generalized predictive approach-based torque and pitch control. *Int. J. Electr. Power Energy Syst.* **2023**, *154*, 109416. [CrossRef]
5. Porté-Agel, F.; Bastankhah, M.; Shamsoddin, S. Wind-turbine and wind-farm flows: A review. *Bound.-Layer Meteorol.* **2020**, *174*, 1–59. [CrossRef] [PubMed]
6. Njiri, J.G.; Söffker, D. State-of-the-art in wind turbine control: Trends and challenges. *Renew. Sustain. Energy Rev.* **2016**, *60*, 377–393. [CrossRef]
7. Rajendran, S.; Diaz, M.; Devi, V.K.; Jena, D.; Travieso, J.C.; Rodriguez, J. Wind Turbine Emulators—A Review. *Processes* **2023**, *11*, 747. [CrossRef]
8. Chiang, M.H. A novel pitch control system for a wind turbine driven by a variable-speed pump-controlled hydraulic servo system. *Mechatronics* **2011**, *21*, 753–761. [CrossRef]
9. Asgharnia, A.; Shahnazi, R.; Jamali, A. Performance and robustness of optimal fractional fuzzy PID controllers for pitch control of a wind turbine using chaotic optimization algorithms. *ISA Trans.* **2018**, *79*, 27–44. [CrossRef]
10. Sierra-Garcia, J.E.; Santos, M.; Pandit, R. Wind turbine pitch reinforcement learning control improved by PID regulator and learning observer. *Eng. Appl. Artif. Intell.* **2022**, *111*, 104769. [CrossRef]
11. Hendradewa, A.P.; Yin, S. Comparative analysis of offshore wind turbine blade maintenance: RL-based and classical strategies for sustainable approach. *Reliab. Eng. Syst. Saf.* **2025**, *253*, 110477. [CrossRef]
12. Xie, J.; Dong, H.; Zhao, X. Data-driven torque and pitch control of wind turbines via reinforcement learning. *Renew. Energy* **2023**, *215*, 118893. [CrossRef]
13. Ma, R.; Siaw, F.L.; Thio, T.H.G. A Hybrid Fuzzy LQR-PI Blade Pitch Control Scheme for Spar-Type Floating Offshore Wind Turbines. *J. Mar. Sci. Eng.* **2024**, *12*, 1306. [CrossRef]
14. Hassan, A.; Ahmad, G.; Shafiullah, M.; Islam, A.; Alam, M.S. Review of the intelligent frameworks for pitch angle control in wind turbines. *IEEE Access* **2025**, *13*, 29864–29885. [CrossRef]
15. Kim, K.; Kim, H.G.; Song, Y.; Paek, I. Design and simulation of an LQR-PI control algorithm for medium wind turbine. *Energies* **2019**, *12*, 2248. [CrossRef]
16. Wen, H.; Liu, Y.; Lee, S.M.; Park, J.H. Optimal power regulating for wind turbine with LQR and disturbance accommodation pitch control. In *Proceedings of the 2023 9th International Conference on Control, Automation and Robotics (ICCAR)*; IEEE: Piscataway, NJ, USA, 2023; pp. 262–267.
17. Engels, W.P.; Van der Hoek, D.; Yu, W.; Reyes Baez, R. Scheduled state feedback control of a wind turbine. *J. Phys. Conf. Ser.* **2020**, *1618*, 022064. [CrossRef]
18. Schlipf, D.; Grau, P.; Raach, S.; Duraiski, R.; Trierweiler, J.; Cheng, P.W. Comparison of linear and nonlinear model predictive control of wind turbines using LIDAR. In *Proceedings of the 2014 American Control Conference*; IEEE: Piscataway, NJ, USA, 2014; pp. 3742–3747.
19. Song, D.; Tu, Y.; Wang, L.; Jin, F.; Li, Z.; Huang, C.; Joo, Y.H. Coordinated optimization on energy capture and torque fluctuation of wind turbines via variable weight NMPC with fuzzy regulator. *Appl. Energy* **2022**, *312*, 118821. [CrossRef]
20. El-Baklish, S.K.; El-Badawy, A.A.; Frison, G.; Diehl, M. Nonlinear model predictive pitch control of aero-elastic wind turbine blades. *Renew. Energy* **2020**, *161*, 777–791. [CrossRef]
21. Lasheen, A.; Saad, M.S.; Emara, H.M.; Elshafei, A.L. Tube-based explicit model predictive output-feedback controller for collective pitching of wind turbines. *Renew. Energy* **2019**, *131*, 549–562. [CrossRef]
22. Fu, D.; Kong, L.; Gong, L.; Wang, A.; Jia, H.; Zhao, N. Wind turbine load optimization control strategy based on LIDAR feed-forward control for primary frequency modulation process with pitch angle reservation. *Energies* **2023**, *16*, 510. [CrossRef]
23. Russell, A.J.; Collu, M.; McDonald, A.S.; Thies, P.R.; Keane, A.; Quayle, A.R. LIDAR-assisted feedforward individual pitch control of a 15 MW floating offshore wind turbine. *Wind Energy* **2024**, *27*, 341–362. [CrossRef]
24. Miquelez-Madariaga, I.; Lizarraga-Zubéldia, I.; de Corcuera, A.D.; Elso, J. Linear uncertain modelling of LIDAR systems for robust wind turbine control design. *Renew. Energy* **2023**, *206*, 367–374. [CrossRef]
25. Yamaguchi, A.; Yousefi, I.; Ishihara, T. Reduction in the fluctuating load on wind turbines by using a combined nacelle acceleration feedback and lidar-based feedforward control. *Energies* **2020**, *13*, 4558. [CrossRef]
26. Han, B.; Zhou, L.; Zhang, Z. LIDAR-assisted radial basis function neural network optimization for wind turbines. *IEEJ Trans. Electr. Electron. Eng.* **2018**, *13*, 195–200. [CrossRef]
27. Bao, J.; Yue, H. Design and assessment of a LIDAR-based model predictive wind turbine control. *Energies* **2022**, *15*, 6429. [CrossRef]
28. Woolcock, L.; Liu, V.; Witherby, A.; Schmid, R.; Mahdizadeh, A. Comparison of REWS and LIDAR-based feedforward control for fatigue load mitigation in wind turbines. *Control Eng. Pract.* **2023**, *138*, 105477. [CrossRef]
29. Jonkman, J.; Butterfield, S.; Musial, W.; Scott, G. *Definition of a 5-MW Reference Wind Turbine for Offshore System Development*; National Renewable Energy Laboratory (NREL): Golden, CO, USA, 2009.
30. *IEC 61400-1:2005*; Wind Turbines Part 1: Design Requirements. International Electrotechnical Commission: Geneva, Switzerland, 2005.

31. *IEC 61400-50-3:2022*; Wind Energy Generation Systems-Part 50-3: Use of Nacelle-Mounted Lidars for Wind Measurements. International Electrotechnical Commission: Geneva, Switzerland, 2022.
32. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal policy optimization algorithms. *arXiv* **2017**, arXiv:1707.06347. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.