

DOI: 10.19666/j.rlfed.202501006

# 基于强化学习的风电机组偏航系统 模型预测控制

桑申刚<sup>1</sup>, 李桂朋<sup>1</sup>, 王向伟<sup>1</sup>, 刘毅<sup>1</sup>, 王森<sup>1</sup>, 申向荣<sup>2</sup>

(1. 华能新能源股份有限公司河北分公司, 河北 石家庄 050011;

2. 华北电力大学自动化系, 河北 保定 071003)

**[摘要]** 提升风电机组偏航系统在多工况场景下的动态性能至关重要。基于此, 提出了一种基于强化学习的风电机组偏航系统模型预测控制策略, 通过双延迟深度确定性策略梯度 (TD3) 算法实现多目标参数动态优化。首先, 针对功率损失率和偏航执行器使用率 2 个相互矛盾的控制目标, 建立偏航系统的多步模型预测控制器 (yaw model predictive control, YMPC); 其次, 针对偏航系统的优化目标及风况特性设计双延迟深度确定性策略梯度 (TD3) 智能体的输入状态、动作和奖励机制, 从而使用 TD3 智能体整定 YMPC 的权重系数和控制步长; 最后, 采用中国北方风电场典型日数据验证了该方法的有效性。结果表明, 与固定控制参数的 YMPC 相比, 所提出的策略显著提升了偏航系统的综合性能。

**[关键词]** 风电机组; 偏航系统; 强化学习; 控制参数

**[引用本文格式]** 桑申刚, 李桂朋, 王向伟, 等. 基于强化学习的风电机组偏航系统模型预测控制[J]. 热力发电, 2025, 54(9): 86-94. SANG Shengang, LI Guipeng, WANG Xiangwei, et al. Predictive control of wind turbine yaw system model based on reinforcement learning[J]. Thermal Power Generation, 2025, 54(9): 86-94.

## Predictive control of wind turbine yaw system model based on reinforcement learning

SANG Shengang<sup>1</sup>, LI Guipeng<sup>1</sup>, WANG Xiangwei<sup>1</sup>, LIU Yi<sup>1</sup>, WANG Sen<sup>1</sup>, SHEN Xiangrong<sup>2</sup>

(1. Huaneng New Energy Co., Ltd. Hebei Branch, Shijiazhuang 050011, China;

2. Department of Automation, North China Electric Power University, Baoding 071003, China)

**Abstract:** It is crucial to improve the dynamic performance of the yaw system of wind turbines in multiple operating scenarios. Therefore, a predictive control strategy for wind turbine yaw system model based on reinforcement learning is proposed, which achieves multi-objective parameter dynamic optimization through the dual-delay depth deterministic policy gradient (TD3) algorithm. Firstly, a multi-step model predictive controller for the yaw system (YMPC) is established to address the conflicting control objectives of power loss rate and yaw actuator utilization rate. Secondly, based on the optimization objectives and wind conditions of the yaw system, a dual-delay depth deterministic strategy gradient (TD3) intelligent agent is designed to determine the input state, action, and reward mechanism of the YMPC. The TD3 intelligent agent is then used to tune the weight coefficients and control step size of the YMPC. Finally, the effectiveness of this method was validated using typical daily data from wind farms in northern China. The results indicate that the proposed strategy significantly improves the overall performance of the yaw system compared with the YMPC with fixed control parameters.

**Key words:** wind turbines; yaw system; reinforcement learning; control parameter

2024 年, 国家能源总局发布关于印发《2024 年 能源工作指导意见》, 指出风力发电、光伏发电是

收稿日期: 2025-01-22 网络首发日期: 2025-06-30

基金项目: 中国华能集团有限公司总部科技项目 (HNKJ22-HF69)

Supported by: Science and Technology Project of China Huanneng Group Co., Ltd. (HNKJ22-HF69)

第一作者简介: 桑申刚 (1982), 男, 工程师, 主要研究方向为电力系统分析、运行与控制, shengang\_sang@hnr.com.cn.

通信作者简介: 申向荣 (1999), 男, 硕士研究生, 主要研究方向为风电机组偏航系统控制、偏航角校正, sxr1984063524@163.com.

发展的重中之重,截至2023年底,全国风电累计装机容量44134万千瓦<sup>[1]</sup>。因此,开发先进的风电机组控制技术以提高其发电效率是当前的研究热点。

风电机组的偏航控制指通过转动机舱位置来保证风机叶片始终正对风向<sup>[2]</sup>。偏航失准不仅会导致能量损失,而且会缩短风电机组的使用寿命。据统计,偏航故障率在风电机组的所有故障中位列前三<sup>[3]</sup>。因此,为了提升风电机组的经济效益,降低机组故障率,需要更加先进的风电机组偏航控制策略。传统偏航控制策略基于机舱顶部风速仪的滤波风向<sup>[4]</sup>。然而,滤波后的风向测量值仅反映过去一段时间的机组偏航误差,偏航动作总是落后于风向的动态变化,同时传统的偏航控制方式的控制精度依赖于风向标的测量精度<sup>[5-6]</sup>。

为了解决风向滞后、测风精度等问题,风向预测开始大量应用于偏航控制策略的研究中。文献[7]结合一维卷积神经网络和长短时记忆(LSTM)神经网络实现对风向的预测并结合风向变化来驱动机组偏航对风。文献[8]依赖激光雷达的精度进行短期风向预测来实现偏航驱动。然而,基于风向预测的偏航控制策略提高了发电效率的同时,却也忽略了风电机组载荷对风机整体寿命的影响。文献[9]提出综合疲劳分析和最优偏航控制成本函数来求出最佳偏航角以此优化目标风机的疲劳载荷,但是也并未充分考虑到机组发电效率问题。

针对上述情况,模型预测控制(MPC)作为一种能够实现多目标优化、可以在有限时间内获得最优控制集的控制策略,已经广泛应用于风电机组控制领域并且取得了较好的控制效果。文献[10]将风向预测模型和一步模型预测控制策略组合用于风电机组,并与工业控制方案相比,所提出的方案能够通过适量的偏航动作来减小偏航误差,提升风电机组发电量。然而,该策略使用了预定义的控制参数。模型预测控制中不同的控制参数对控制方法的潜在性能影响非常重要。如文献[11]提出了一种利用动态权重因子的自适应模型预测控制方法,该方法提高了变工况下系统的整体性能。文献[12]提出了一种基于模糊推导权重系数评估器的偏航模型预测控制方法,在改变权重因子的同时优化模糊规则和隶属度函数。文献[13]使用贝叶斯算法优化了MPC控制器的预测时域和权重系数。但模糊规则和贝叶斯算法都依赖专家知识和经验构造模型

或规则,在应对模型不确定性和动态变化方面适应性较差,难以保证全局最优。因此需要一种不依赖于专家的先验知识和规则,能够根据环境变化自适应地优化控制参数的方法。

无模型强化学习算法为突破上述局限提供了新途径<sup>[14-16]</sup>。文献[17]提出基于强化学习的风电机组转矩增益系数控制策略提高了机组在湍流风下的平均风能捕获效率。文献[18]提出的一种基于改进深度强化学习的风电场尾流优化控制策略在不同风速下均能有效优化尾流分布,可提升风电场整体功率最大至34%。文献[19]提出的无模型强化学习俯仰角控制器训练6个智能体来输出俯仰角控制输出,确保了控制器的稳定性。

因此,将强化学习模型应用于风电机组控制领域值得研究。目前一些学者通过对MPC控制参数的整定,极大改善了控制任务的性能,然而,鲜有对控制步长进行研究,MPC的控制步长是影响控制性能的关键参数,较短的控制步长可以提高控制的灵敏度和精度,但也会增加计算量和实时性要求;而较长的控制步长则可以降低计算负担,但可能导致控制性能下降<sup>[20]</sup>。

针对上述问题,本文提出了一种基于强化学习整定多步模型预测控制器(yaw model predictive control, YMPC)参数的策略,首先建立YMPC,然后设计强化学习智能体的输入状态、动作和奖励机制,利用智能体在变工况情况下实时调整权重系数和控制步长,通过控制参数的实时优化来提高对风精度,减少偏航执行器的使用次数。

## 1 水平轴风电机组多步模型预测控制

### 1.1 水平轴风机的偏航系统

本文所涉及的风机类型为双馈异步式,理论上风电功率 $P_{\text{ext}}$ 可以表示为<sup>[21]</sup>:

$$P_{\text{ext}} = \rho A_r C_p(\omega_r, V_o, \beta) V_e^3 / 2 \quad (1)$$

$$V_e = V_o \cos(e) = V_o \cos(\omega - \theta) \quad (2)$$

式中: $\rho$ 为空气密度; $A_r$ 为转子面积; $C_p(\omega_r, V_o, \beta)$ 为与转子转速 $\omega_r$ 、切入流风速 $V_o$ 、桨距角 $\beta$ 相关的气动功率系数; $V_e$ 为垂直于转子平面的有效风速; $e$ 为风向 $\omega$ 与机舱位置 $\theta$ 之间的偏航误差。

从式(1)、式(2)可以得出,最小化偏航误差 $e$ 对风电机组的最大功率提取起着重要作用。风电机组的偏航控制系统结合所得风向的数据,使用偏航执行器调整机舱朝向,通过减小偏航误差 $e$ ,使风轮

始终面对风向，最大化风能捕获效率。

## 1.2 多步偏航模型预测控制系统

本文采用具有有限控制集的 YMPC，YMPC 通过跟踪风向来最大化发电量，同时避免偏航执行器的过度使用，所提出的 YMPC 的控制框图如图 1 所示，主要涉及多步预测模型、目标函数、有限控制集的构建以及 YMPC 的求解方法<sup>[22]</sup>，该方法的具体内容如下。

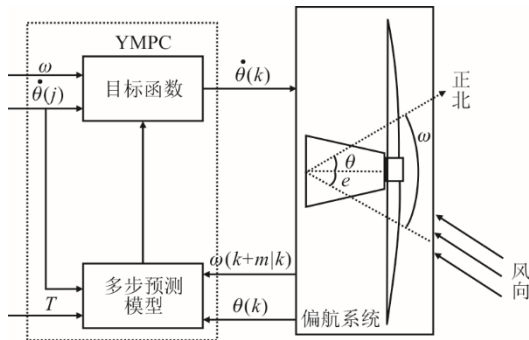


图 1 YMPC 控制框图

Fig.1 The YMPC control block diagram

### 1.2.1 多步预测模型

通过最小化偏航误差  $e$  来提升发电功率，选取偏航误差  $e$  作为状态变量，其一步预测模型为：

$$e(k+1|k) = \omega(k+1|k) - \theta(k+1|k) \quad (3)$$

式中： $k$  为第  $k$  个采样周期； $e(k+1|k)$ 、 $\omega(k+1|k)$  和  $\theta(k+1|k)$  分别为一步偏航误差、一步风向和一步机舱位置的预测值。

机舱通过偏航控制系统以一定的偏航速率旋转，因此机舱位置预测值  $\theta(k+1|k)$  为：

$$\theta(k+1|k) = \theta(k) + \dot{\theta}(k+1)T \quad (4)$$

式中： $\theta(k)$  为第  $k$  个采用周期的机舱位置； $T$  为控制步长； $\dot{\theta}(k+1)$  为第  $k+1$  个采样周期的偏航率。

由式(3)和式(4)得出偏航误差  $e(k+m|k)$  的  $m$  步预测模型为：

$$e(k+m|k) = \omega(k+m|k) - \theta(k) - (\dot{\theta}(k+m) + \dots + \dot{\theta}(k+1))T \quad (5)$$

### 1.2.2 目标函数

由式(1)可知，风电机组的风力提取与偏航误差具有余弦平方关系，所以使用  $\alpha$  来评估功率损失率，可表示为：

$$\alpha = (E_{\text{ideal}} - E_c) / E_{\text{ideal}} = 1 - \left( \sum_{N=k+1}^{k+m} \cos^2(e(N)) \right) / m \quad (6)$$

式中： $E_{\text{ideal}}$  为在  $m$  步预测时域内不考虑偏航误差引

起的功率损失的情况下捕获的理想风能； $E_c$  为实际所捕获到的风能，可表示为：

$$E_c = \sum_{N=k+1}^{k+m} \rho A_r C_p(\omega_r, V_o, \beta) V_o^3 \cos^2(e(N)) / 2 \quad (7)$$

偏航执行器使用率  $\gamma$  可表示为：

$$\gamma = c_{\text{yaw}} / c_{\text{total}} = \left( \sum_{N=k+1}^{k+m} (|\dot{\theta}(N)| > 0) \right) / m \quad (8)$$

式中： $c_{\text{yaw}}$  为评估时间  $c_{\text{total}}$  内偏航执行器动作时间。

在式(6)一式(8)的基础上，通过权重系数  $\omega \in [0, 1]$ ，将  $\alpha$  和  $\gamma$  2 个控制器目标组合起来。因此， $m$  步预测范围内的目标函数  $Q_F$  表达为：

$$Q_F = (1 - \omega) \left( 1 - \left( \sum_{N=k+1}^{k+m} \cos^2(e(N)) \right) / m \right) + \omega \left( \sum_{N=k+1}^{k+m} (|\dot{\theta}(N)| > 0) \right) / m \quad (9)$$

### 1.2.3 有限控制集和约束

研究所涉及 YMPC 控制变量为横摆角速度，有 3 个可能的元素：

$$\dot{\theta}(j) = \begin{cases} -y_{\text{aw\_rate}}, & j = 0 \\ y_{\text{aw\_rate}}, & j = 1 \\ 0, & j = 2 \end{cases} \quad (10)$$

式中： $\dot{\theta}(j)$  为与  $j$  有关的偏航率。

此外，因为本文所涉及的风电机组偏航系统是刚性系统，所以偏航率要受到如下约束：

$$\dot{\theta}(k+1) = \begin{cases} \dot{\theta}(k+1)\{j=0,1,2\}, \exists \dot{\theta}(k)\{j=0\} \\ \dot{\theta}(k+1)\{j=0,1\}, \exists \dot{\theta}(k)\{j=1\} \\ \dot{\theta}(k+1)\{j=0,2\}, \exists \dot{\theta}(k)\{j=2\} \end{cases} \quad (11)$$

根据式(9)一式(11)提出的 YMPC 方法，需要求解横摆角速度的控制序列，而控制序列的数量取决于预测时域  $m$  的大小，通过文献[23]中提出的穷举搜索方法可以实现求解  $Q_F$  最小值的控制序列，该控制序列中的第一个动作作为下一控制周期的偏航动作。

## 2 基于 TD3 的 YMPC 控制参数整定

本文提出的基于双延迟深度确定性策略梯度的偏航系统多步模型预测控制 (YMPC-TD3) 总体控制框图如图 2 所示。

### 2.1 TD3 算法结构

强化学习是一种典型的经验驱动、自主学习的方法，它通过智能体与环境的互动来探寻最佳行为策略<sup>[24]</sup>。其中 TD3 算法是目前较为先进的强化学习

算法之一。此算法的强化学习过程为：智能体感知环境状态 (states)，依据既定策略 (policy) 采取行动 (actions)，从而获取奖励 (reward) 以改善智能体自身表现<sup>[24]</sup>，其整体网络结构如图 3 所示。

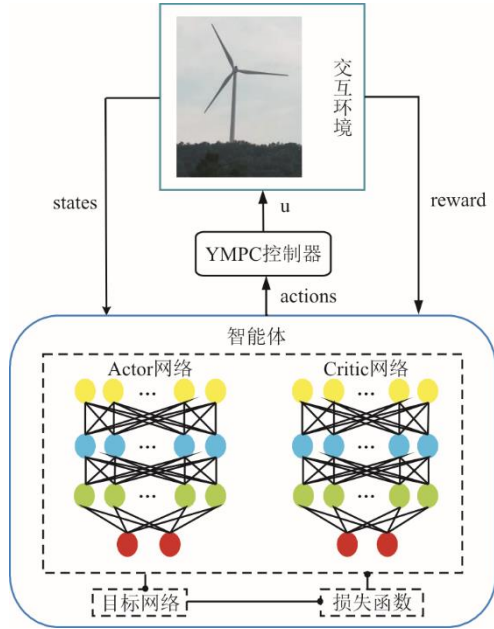


图 2 YMPC-TD3 控制框图  
Fig.2 The YMPC-TD3 control block diagram

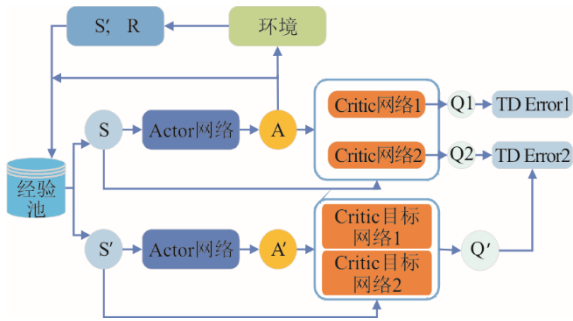


图 3 TD3 整体网络结构  
Fig.3 The overall network structural diagram of TD3

TD3 算法由 Actor 网络和 Actor 目标策略网络、2 个独立的 Critic 网络及 Critic 目标网络 6 个网络组成。Actor 网络负责构建状态到动作的映射，为智能体在各种状态下选择行动提供指导<sup>[25]</sup>。Actor 网络的更新通过确定性策略梯度算法实现：

$$\nabla_{\phi} J(\phi) = E_{s \sim p_{\pi}} \left[ \nabla_a Q^{\pi}(s, a) \Big|_{a=\pi(s)} \nabla_{\phi} \pi(s) \right] \quad (12)$$

$$Q_{\pi}(s, a) = E_{s \sim p_{\pi}, a \sim \pi} [R_t | s, a] \quad (13)$$

式中： $\nabla$  为权重参数的梯度； $Q^{\pi}(s, a)$  为状态  $s$  下执行动作  $a$  的  $Q$  函数； $E_{s \sim p_{\pi}}$  为状态  $s$  下采用策略  $\pi$  所作出动作时的预期收益。

Critic 网络用于建立状态到价值的映射，以评

估 Actor 网络建立映射的价值，被称为价值  $Q$  函数。在传统的强化学习模型中，价值网络常常高估由 Actor 网络生成动作的价值，即其估计的  $Q$  值往往高于实际的  $Q$  值<sup>[26]</sup>。为解决此问题，TD3 采用了双重 Critic 网络，分别对应 2 个  $Q$  函数，记作  $Q_1(s, a)$  和  $Q_2(s, a)$ ，取两者的最小值作为 Critic 网络的结果。此外，Critic 网络和 Critic 目标策略网络的评估函数基于贝尔曼方程进行延迟学习<sup>[27]</sup>：

$$Q^{\pi}(s, a) = r + \gamma E_{s', a'} \left[ Q^{\pi}(s', a') \right], a' \sim \pi(s') \quad (14)$$

式中： $(s', a')$  为下一时刻的状态和动作。

此外，还加入了目标策略噪声抑制确定性策略中出现的估值函数过拟合问题<sup>[28]</sup>：

$$\gamma = r + \gamma \min(Q_{01}, Q_{02})(s', \pi_{\phi}(s') + \varepsilon) \quad (15)$$

$$\varepsilon = \text{clip}(N(0, \sigma), -c, c) \quad (16)$$

式中： $\varepsilon$  为高斯噪声； $c$  为策略平滑噪声的截断边界值；clip 为截断函数。

同时，Actor 网络的更新频率应该低于评估网络，因此采用一种软更新策略<sup>[29]</sup>：

$$\theta_i = \tau \theta_i + (1 - \tau) \theta_i \quad (17)$$

$$\phi_i = \tau \phi_i + (1 - \tau) \phi_i \quad (18)$$

式中： $\tau$  为软更新率。

## 2.2 YMPC-TD3 控制器设计

### 2.2.1 动作空间

动作空间即智能体所采取的动作，本文需要使用 TD3 整定的参数为控制步长  $T$  和权重系数  $\omega$ ，控制步长的选取需要限定在一定范围内，控制步长过长或者过短都会影响系统的灵敏度和稳定性，同时需要控制权重系数  $\omega$  在一定范围内以防偏航系统锁死，所以设计动作空间  $a$  为：

$$\begin{cases} a = [\omega \ T] \\ \omega_l < \omega < \omega_h \\ T_l \leq T \leq T_h \end{cases} \quad (19)$$

式中： $\omega_l$ 、 $\omega_h$  分别为  $\omega$  的下限值和上限值； $T_l$ 、 $T_h$  分别为  $T$  的下限值和上限值。

### 2.2.2 状态空间

状态空间是强化学习与环境交互的核心，状态空间包括 YMPC 在控制参数整定时所需的信息，状态空间选取为：

$$s_t = [e(k), \text{Inc}_m, \text{Rt}_m] \quad (20)$$

式中： $e(k)$  为风机当前时间步的偏航误差； $\text{Inc}_m$ 、 $\text{Rt}_m$  分别为预测范围  $m$  步内的预测风向的增量和预测风向的变化率。

其中  $\text{Inc}_m$ 、 $\text{Rt}_m$  计算公式分别为:

$$\text{Inc}_m = \text{ABS}(\omega(k+m|k) - \theta(k+1|k)) \quad (21)$$

$$\text{Rt}_m = \sqrt{\sum_{i=1}^m (\omega(k+i|k) - \text{mean}(\omega))^2 / m} \quad (22)$$

式中:  $\text{ABS}$  表示绝对值;  $\text{mean}$  表示预测风向的平均值。

### 2.2.3 奖励机制

奖励函数的设计是强化学习策略不断进步的基础,设计合理的奖励对 YMPC 的参数整定起着引导性的作用,奖励函数包括正奖励和负奖励,正奖励鼓励智能体更快完成学习任务,负奖励避免智能体做出不利于系统的行为<sup>[29]</sup>。

本文所设计的奖励函数  $R$  由目标函数奖励项  $r_1$ , 权重系数变化率奖励项  $r_2$  和控制步长变化率奖励项  $r_3$  组成:

$$R = r_1 + r_2 + r_3 \quad (23)$$

奖励函数的设计是为了让机组偏航系统适应不同的工况且使得质量函数  $Q_F$  最小化,故有:

$$r_1 = \begin{cases} k_1(1 - \lg Q_F), & 0 < Q_F < 0.1 \\ -200, & 0.1 \leq Q_F \end{cases} \quad (24)$$

式中:  $k_1$  为  $r_1$  奖励项的权重。

由于参数的变动会导致系统稳定性影响,因此对权重系数  $\omega$  的一阶导数和控制步长  $T$  的一阶导数设置了惩罚,期望强化学习策略不断进步的基础上,保持参数变化的稳定性,奖励  $r_2$ 、 $r_3$  分别为:

$$r_2 = k_2 \left(1 - \frac{|\dot{\omega}|}{\dot{\omega}_0}\right) \quad (25)$$

$$r_3 = k_3 \left(1 - \frac{|\dot{T}|}{\dot{T}_0}\right) \quad (26)$$

式中:  $\dot{\omega}_0$ 、 $\dot{T}_0$  分别为权重系数  $\omega$  和控制步长  $T$  的一阶导数阈值;  $k_2$ 、 $k_3$  分别为  $r_2$  和  $r_3$  奖励项的权重。

## 3 算法验证

选取河北某风电场的典型日的风向数据,风向状况会影响偏航控制系统的性能,本文选取了 3 个典型日的风向数据,如图 4 所示。

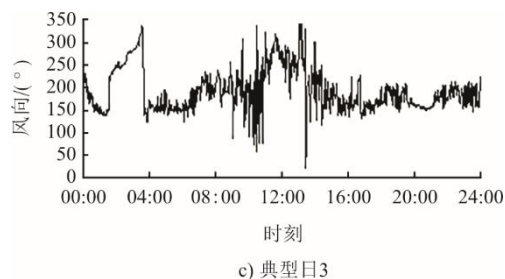
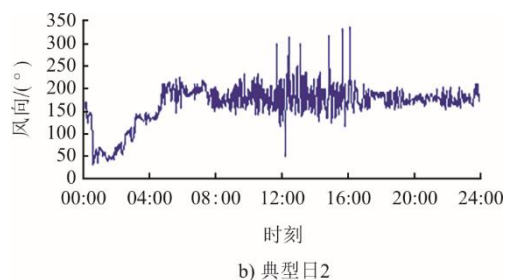
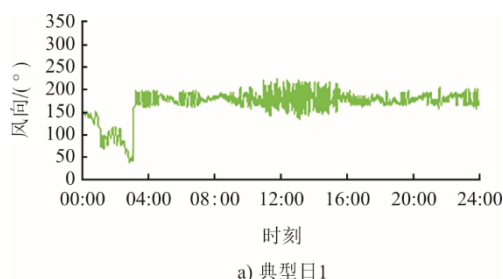


图 4 典型日风向数据

Fig.4 The wind direction data on typical days

量化风向变化状况的风向变化量  $\delta(\omega)$  为:

$$\delta(\omega) = \sqrt{\frac{\sum_{k=2}^N (\omega(k) - \omega(k-1))^2}{(N-1)}}, N = 86\ 400 \quad (27)$$

本文使用具有不同控制步长的 3 个 YMPC 与所提出的 YMPC-TD3 作对比。偏航率设置为  $0.5^\circ/\text{s}$ , 权重系数  $\omega$  设置为 0.015, 3 个 YMPC 的控制步长参数分别为  $T=4$ 、5、6, 所提出的 YMPC 建立在 MATLAB 软件上,并且在内存为 4 GB、CPU 为 1.6 GHz 的笔记本电脑上运行。YMPC-TD3 主要超参数见表 1。

表 1 YMPC-TD3 主要超参数

Tab.1 The main hyperparameters of YMPC-TD3

项目	数值
折扣因子	0.99
经验池大小	106
经验采样数	64
Actor 网络学习率	0.000 1
Critic 网络学习率	0.001
软更新系数	0.005
延迟更新参数	2

### 3.1 预训练

为确保偏航动作的持续性与系统动态响应相匹配,结合偏航执行器约 2 s 的启动延迟特性,控制步长需满足  $T > T_{\text{start}}$  ( $T_{\text{start}}=2$  s)。为避免步长过长引发控制精度衰减,同时抑制步长过短导致的频繁动作,基于系统稳定性与灵敏度的权衡分析,设定

控制步长参数范围为  $T \in [4, 10]$  s。此外，为避免权重系数极端值引起的控制目标失衡，动作空间中的权重系数  $\omega$  被约束为  $\omega \in [0, 0.15]$ 。奖励函数中  $k_1$ 、 $k_2$  和  $k_3$  分别设置为 20、100 和 100。训练轮次为 400 轮，每轮训练 3 000 步，实际的训练过程奖励如图 5 所示。由图 5 可以看到，在前 150 个回合，智能体通过不断尝试使得奖励不断上升，当训练轮次达到 200 次以后，奖励趋于收敛。

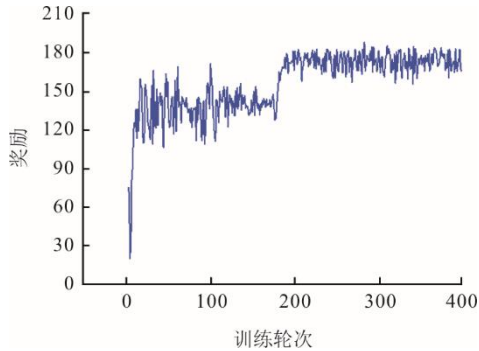


图 5 YMPC-TD3 训练过程奖励

Fig.5 Rewards during the YMPC-TD3 training process

### 3.2 性能比较

为了评估本文提出的控制策略的性能，需选取 3 个典型日（风向变化量分别为 3.64、5.76 和 7.15）进行对比分析。首先基于典型日 1 数据对 YMPC( $T=4$ )、YMPC( $T=5$ )、YMPC( $T=6$ ) 及 YMPC-TD3 进行性能验证，随后扩展至典型日 2、典型日 3 考察多风况下的综合表现。

#### 3.2.1 控制器性能比较

图 6 显示了典型日 1 下 YMPC( $T=4$ )、YMPC( $T=5$ )、YMPC( $T=6$ ) 和 YMPC-TD3 的仿真结果。将 4 个控制器的结果集中起来，通过偏航误差以及 1.2.2 节目标函数中的偏航执行器使用率和功率损失率来评估控制器的性能。图 6 显示 4 个 YMPC 控制器都在一定程度上实现了风向跟踪，4 个控制器都处于正常工作状态。然而，不同控制器下的偏航误差、偏航执行器使用率和功率损失率存在一定的差异。偏航误差可根据 1.2.2 小节中的式(6)来衡量，因此此处仅考虑功率损失率和偏航执行器使用率。图 6c)和图 6d)显示 YMPC( $T=4$ ) 的偏航致动器使用率 (9.63%) 排在第 3 位，处于次优水平，同时它的功率损失达到最优值 (1.79%)，而 YMPC-TD3 损失的功率与 YMPC( $T=5$ )相近，相比 YMPC( $T=5$ )增加了 0.03 百分点，但是其偏航致动器使用率最小 (8.34%)，综合性能良好。

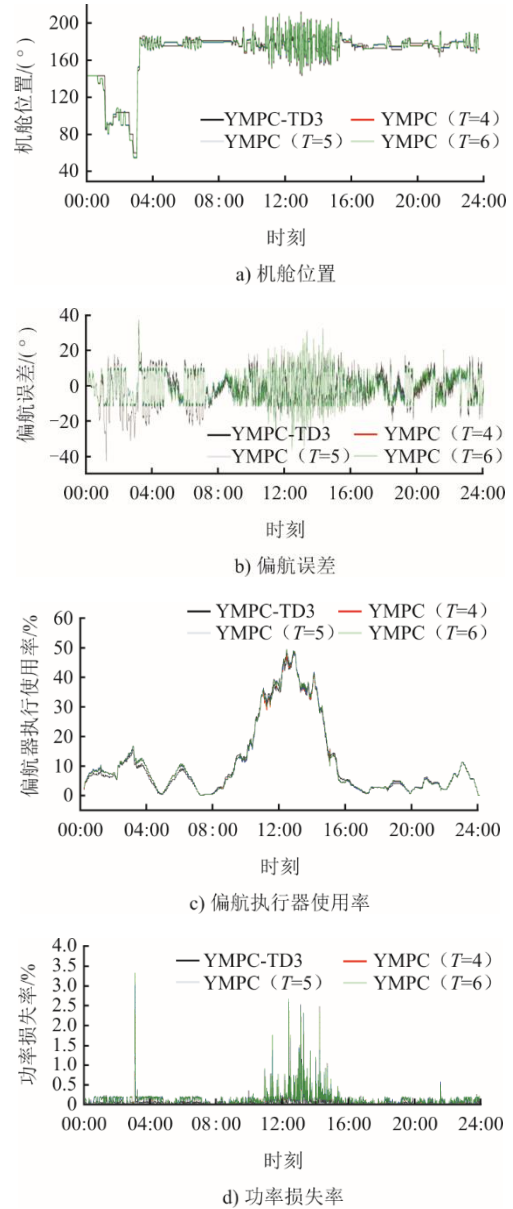


图 6 控制器仿真结果

Fig.6 Simulation results of the controllers

图 7 展示了 YMPC-TD3 策略中权重系数动态调整对系统性能的影响。

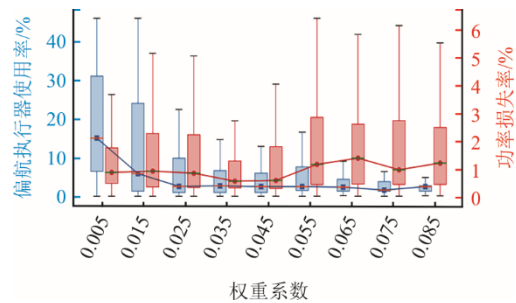


图 7 YMPC-TD3 不同权重系数下偏航执行器使用率和功率损失率趋势

Fig.7 Trends of yaw actuator utilization and power loss rate with different weight coefficients of YMPC-TD3

由图 7 仿真结果表明, 功率损失率虽呈现小幅度波动, 但是整体呈现上升的趋势(增幅约 0.6%); 而偏航执行器使用率的变化趋势较为明显, 随着权重系数的增加, 偏航执行器使用率呈现下降的趋势(降幅达 12%)。这些变化可以由式(6)一式(9)直观解释, 随着  $\omega$  的增加,  $\alpha$  在目标函数  $Q_F$  中占比降低, YMPC 在计算最佳控制动作时将更侧重于考虑降低偏航执行器使用率, 从而产生权重系数与偏航执行器使用率负相关、与功率损失率正相关的现象。

为了进一步比较结果, 将典型日 1 下 YMPC ( $T=4$ )、YMPC ( $T=5$ )、YMPC ( $T=6$ ) 和 YMPC-TD3 的平均偏航执行器使用率和平均功率损失率的具体数值汇总于表 2。由表 2 可见, 在 4 种 YMPC 中, 本文提出的 YMPC-TD3 控制器的综合性能最好。

表 2 典型日 1 数值统计 单位: %  
Tab.2 Numerical statistics of typical day 1

项目	平均偏航执行器使用率	平均功率损失率
YMPC-TD3	8.34	1.85
YMPC ( $T=4$ )	9.63	1.79
YMPC ( $T=5$ )	8.81	1.82
YMPC ( $T=6$ )	10.32	2.17

### 3.2.2 不同风况敏感性分析

为探究不同风向动态特性对控制策略的影响, 本文进一步分析了不同风向变化量下各控制器的性能。表 3 和表 4 显示了典型日 2 和典型日 3 时 4 种 YMPC 控制器的仿真结果。

表 3 典型日 2 数值统计 单位: %  
Tab.3 Numerical statistics of typical day 2

项目	平均偏航执行器使用率	平均功率损失率
YMPC-TD3	12.04	1.68
YMPC ( $T=4$ )	13.14	1.97
YMPC ( $T=5$ )	11.83	1.91
YMPC ( $T=6$ )	11.78	2.14

表 4 典型日 3 数值统计 单位: %  
Tab.4 Numerical statistics of typical day 3

项目	平均偏航执行器使用率	平均功率损失率
YMPC-TD3	13.28	1.72
YMPC ( $T=4$ )	12.68	2.37
YMPC ( $T=5$ )	13.60	2.13
YMPC ( $T=6$ )	14.95	1.86

由表 3 和表 4 可以看到, YMPC-TD3 的综合性能(偏航执行器使用率 12.04% 和 13.28%, 功率损

失率 1.68% 和 1.72%) 仍然优于 YMPC。还可以观察到, 风向变化量越大, 偏航执行器使用率会随之增大, 而功率损失率在不同 YMPC 控制器下有所不同。本文以 YMPC-TD3 和 YMPC ( $T=5$ ) 的控制结果为例进行分析。

图 8 展示了 YMPC ( $T=5$ ) 和 YMPC-TD3 在 3 个典型日下的偏航执行器使用率。由图 8 可见, 随着风向状况复杂, 偏航执行器使用率会随之增加。

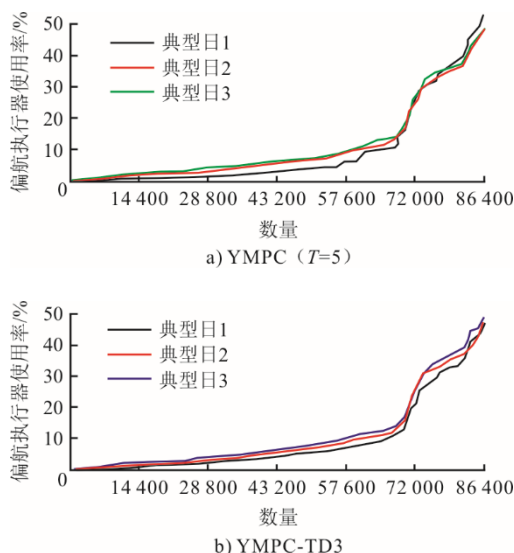
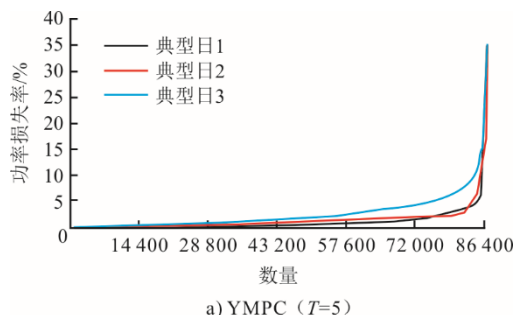
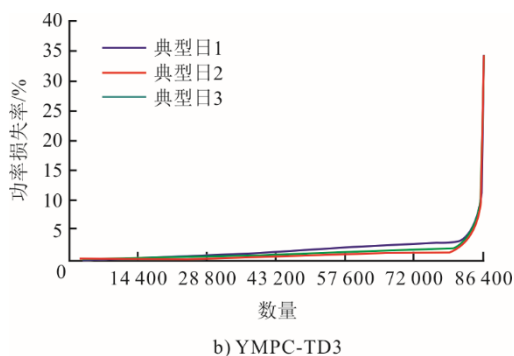


图 8 偏航执行器使用率  
Fig.8 Utilization rates of the yaw actuators

图 9 为 3 个典型日的功率损失率。由图 9 可见, YMPC ( $T=5$ ) 在典型日 1 和典型日 2 的功率损失率相差不大, 而典型日 3 的偏航执行率在其中最大, 相比较于典型日 1 和典型日 2 下分别增加了 4.79 个百分点和 1.77 个百分点, 但是功率损失率也最大, 相较于典型日 1 增加了 0.31 个百分点。YMPC-TD3 在典型日 1 的功率损失率最大, 在典型日 2 的功率损失率最小(相对降低 0.17 百分点), 典型日 3 居中(相对降低 0.13 百分点)。结果表明, 本文提出的 YMPC-TD3 在风况变化的情况下, 依旧能维持其基本性能, 改善普通的 YMPC 在风向情况复杂条件下无法减小功率损失的问题, 减少了无效偏航。





b) YMPC-TD3  
图9 功率损失率  
Fig.9 Power loss rate

## 4 结 论

为了提升偏航系统综合性能, 应对风况变化, 本文提出了一种基于强化学习的风电机组偏航系统模型预测控制策略, 该方法通过 TD3 算法对 YMPC 的控制参数进行整定。YMPC 的性能受到机组工况的影响, 因此设置机组运行工况信息作为智能体输入的状态, 以目标函数  $Q_F$  最小化和参数稳定性为目标制定奖励函数, 通过智能体输出的动作实时调整 YMPC 的权重系数和控制步长, 并且利用典型日数据, 仿真验证了所提方法的有效性。

1) 通过 TD3 算法整定 YMPC 的控制参数, 并利用典型日进行仿真, 结果表明, 与普通的 YMPC 相比, 本文所提出的 YMPC-TD3 综合性能更佳。

2) 在动态风向条件下, 普通 YMPC 策略虽提高了偏航执行器动作频率, 却未能有效降低风机功率损耗。本文提出的 YMPC-TD3 算法通过优化控制参数, 在维持偏航执行器执行器合理使用率的同时, 与典型日 1 相比分别将功率损耗降低了 0.17 百分点和 0.13 百分点, 实现了复杂风向状况下偏航系统综合性能的提升。

3) 本文所提出的 TD3 参数整定仅仅是针对权重系数和控制步长进行整定, 而对预测时域整定的问题值得进一步研究。

### 【参 考 文 献】

[1] 国家能源局. 2024 年能源工作指导意见[EB/OL]. (2024-04-01)[2025-01-10]. <https://www.nea.gov.cn>. National Energy Administration. Guiding opinions on energy work in 2024[EB/OL]. (2024-04-01)[2025-01-10]. <https://www.nea.gov.cn>.

[2] 肖玥, 封焯文, 余子鹏, 等. 基于风向预测的新型偏航控制方法[J]. 太阳能学报, 2021, 42(2): 144-149. XIAO Yue, FENG Zhuowen, YU Zipeng, et al. A novel yaw control method for wind turbines based on predicted wind directions[J]. Acta Energetica Solaris Sinica, 2021, 42(2): 144-149.

[3] PÉREZ P M J, MÁRQUEZ G P F, TOBIAS A, et al. Wind turbine reliability analysis[J]. Renewable and Sustainable Energy Reviews, 2013, 23: 463-472.

[4] 沈小军, 杜万里. 大型风力发电机偏航系统控制策略研究现状及展望[J]. 电工技术学报, 2015, 30(10): 196-203. SHEN Xiaojun, DU Wanli. Expectation and review of control strategy of large wind turbine yaw system[J]. Transactions of China Electrotechnical Society, 2015, 30(10): 196-203.

[5] 曹弘, 孙建平. 风力发电机偏航系统控制策略研究[J]. 电力设备管理, 2019(2): 81-83. CAO Hong, SUN Jianping. Expectation and review of control strategy of large wind turbine yaw system[J]. Electric Power Equipment Management, 2019(2): 81-83.

[6] 谭芝, 陈众, 汤敏, 等. 风力发电机组偏航优化控制方法[J]. 电力学报, 2014, 29(1): 66-69. TAN Zhi, CHEN Zhong, TANG Min, et al. Wind turbine yaw optimal control method[J]. Journal of Electric Power, 2014, 29(1): 66-69.

[7] 叶锋, 林周泉, 李宁波, 等. 一种基于 CNN-LSTM 风向预测的风电机组偏航控制方法[J]. 电气时代, 2022(9): 43-46. YE Feng, LIN Zhouquan, LI Ningbo, et al. A wind turbine yaw control method based on CNN-LSTM wind direction prediction[J]. Electric Age, 2022(9): 43-46.

[8] CHEN W, LIU H, LIN Y, et al. LSTM-NN yaw control of wind turbines based on upstream wind information[J]. Energies, 2020, 13(6): 1482.

[9] RUIYANG H, HONGXING Y, LIN L. Optimal yaw strategy and fatigue analysis of wind turbines under the combined effects of wake and yaw control[J]. Applied Energy, 2023, 337: 223.

[10] SONG D, FAN X, YANG J, et al. Power extraction efficiency optimization of horizontal-axis wind turbines through optimizing control parameters of yaw control systems using an intelligent method[J]. Applied Energy, 2018, 224: 267-279.

[11] ZHANG S, HUANG L, SONG D, et al. Model predictive yaw control using fuzzy-deduced weighting factor for large-scale wind turbines[J]. Energy Engineering, 2021, 118(2): 237-250.

[12] SONG D, LI Z, DENG X, et al. Deep optimization of model predictive control performance for wind turbine yaw system based on intelligent fuzzy deduction[J]. Expert Systems with Applications, 2023, 221: 119705.

[13] 吴西涛, 魏超, 翟建坤, 等. 考虑横摆稳定性的无人车轨迹跟踪控制优化研究[J]. 机械工程学报, 2022, 58(6): 130-142. WU Xitao, WEI Chao, ZHAI Jiankun, et al. Study on the optimization of autonomous vehicle on path-following considering yaw stability[J]. Chinese Journal of Mechanical Engineering, 2022, 58(6): 130-142.

[14] 高阳, 陈世福, 陆鑫. 强化学习研究综述[J]. 自动化学报, 2004, 30(1): 86-100. GAO Yang, CHEN Shifu, LU Xin. Research on reinforcement learning technology: a review[J]. Journal of Automatica Sinica, 2004, 30(1): 86-100.

[15] 柯恺宸, 金士博, 高博扬, 等. 基于强化学习的机器人多接触交互任务控制[J]. 动力学与控制学报, 2023, 21(12): 53-69. KE Kaichen, JIN Shibo, GAO Boyang, et al. A survey of robot intelligent control method in contact-rich tasks[J]. Journal of Dynamics and Control, 2023, 21(12): 53-69.

- [16] 陈帅. 基于强化学习的微电网能量管理与调度[D]. 北京: 北京科技大学, 2024: 1.  
CHEN Shuai. Microgrid energy management and dispatching based on reinforcement learning[D]. Beijing: University of Science and Technology Beijing, 2024: 1.
- [17] 王宇. 基于强化学习的风电机组转矩控制优化策略研究[D]. 沈阳: 沈阳工业大学, 2024: 1.  
WANG Yu. Research on optimization strategy of wind turbine torque control based on reinforcement learning[D]. Shenyang: Shenyang University of Technology, 2024: 1.
- [18] 张勇, 刘春, 褚梦珂, 等. 计及尾流的改进深度确定性策略梯度风电场功率优化控制策略[J]. 电力系统及其自动化学报, 2025, 37(2): 68-77.  
ZHANG Yong, LIU Chun, CHU Mengke, et al. Improved deep deterministic policy gradient for wind farm power optimal control strategies accounting for wake effect[J]. Proceedings of the CSU-EPSA, 2025, 37(2): 68-77.
- [19] NABEEL A, LASHEEN A, ELSHAFEI L A, et al. Fuzzy-based collective pitch control for wind turbine via deep reinforcement learning[J]. ISA Transactions, 2024, 148: 307-325.
- [20] 黄益绍, 王博, 李晨艳, 等. 自适应 MPC 智能车轨迹跟踪控制[J]. 中国测试, 2025, 51(2): 169-175.  
HUANG Yishao, WANG Bo, LI Chenyan, et al. Intelligent vehicle trajectory tracking control based on adaptive MPC[J]. China Measurement & Test, 2025, 51(2): 169-175.
- [21] 胡兵, 海玲, 程丽娟, 等. 小型风力发电机偏航控制系统实验研究[J]. 微电机, 2020, 53(6): 64-67.  
HU Bing, HAI Ling, CHENG Lijuan, et al. Experimental study on yaw control system of small wind turbine[J]. Micromotor, 2020, 53(6): 64-67.
- [22] SONG D, LI Z, WANG L, et al. Energy capture efficiency enhancement of wind turbines via stochastic model predictive yaw control based on intelligent scenarios generation[J]. Applied Energy, 2022, 312: 118773.
- [23] SONG D R, LI Q A, CAI Z L, et al. Model predictive control using multi-step prediction model for electrical yaw system of horizontal-axis wind turbines[J]. IEEE Transactions on Sustainable Energy, 2019, 10(4): 2084-2093.
- [24] 顾雪平, 刘彤, 李少岩, 等. 基于改进双延迟深度确定性策略梯度算法的电网有功安全校正控制[J]. 电工技术学报, 2023, 38(8): 2162-2177.  
GU Xueping, LIU Tong, LI Shaoyan, et al. Active power correction control of power grid based on improved twin delayed deep deterministic policy gradient algorithm[J]. Transactions of China Electrotechnical Society, 2023, 38(8): 2162-2177.
- [25] 李志军, 徐博, 张家安, 等. 基于 TD3 可变长度时间窗口最优加权的短期负荷预测策略[J]. 电力建设, 2024, 45(6): 140-148.  
LI Zhijun, XU Bo, ZHANG Jiaan, et al. Short-term load optimal weighted forecasting strategy based on TD3 variable length time window[J]. Electric Power Construction, 2024, 45(6): 140-148.
- [26] 管超骏, 雷正玲, 霍海波, 等. 基于强化学习的固体氧化物燃料电池输出电压自抗扰控制研究[J]. 发电技术, 2024, 45(6): 1163-1172.  
GUAN Chaojun, LEI Zhengling, HUO Haibo, et al. Study on active disturbance rejection control of output voltage of solid oxide fuel cell based on reinforcement learning[J]. Power Generation Technology, 2024, 45(6): 1163-1172.
- [27] 陈佐明, 詹捷宇. 基于 TD3 算法的自动协商策略[J]. 计算机系统应用, 2023, 32(3): 15-24.  
CHEN Zuoming, ZHAN Jieyu. Automatic negotiation strategy based on TD3 algorithm[J]. Computer Systems & Applications, 2023, 32(3): 15-24.
- [28] 周娴玮, 包明豪, 叶鑫, 等. 带 Q 网络过滤的两阶段 TD3 深度强化学习方法[J]. 计算机技术与发展, 2023, 33(10): 101-108.  
ZHOU Xianwei, BAO Minghao, YE Xin, et al. Two-stage TD3 deep reinforcement learning algorithm with Q network filtration[J]. Computer Technology and Development, 2023, 33(10): 101-108.
- [29] 喻磊, 谈竹奎, 王扬, 等. 基于深度强化学习的电动汽车有序充电优化方法[J]. 南方电网技术, 2024, 18(12): 148-155.  
YU Lei, TAN Zhukui, WANG Yang, et al. Deep reinforcement learning based optimization method for ordered charging of electric vehicle[J]. Southern Power System Technology, 2024, 18(12): 148-155.

(责任编辑 杜亚勤)