XX Simposio CEA de Control Inteligente
25-27 de junio de 2025, Huelva

Control del ángulo de palas de aerogeneradores basado en aprendizaje por refuerzo y PID

Sierra García, Jesús Enrique[a*], Santos Peñas, Matilde[b]

[a] *Departamento de Digitalización, Universidad de Burgos, Campus Río Vena, Avda. Cantabria, s/n, 09006 Burgos, España.*
[b] *Instituto de Tecnología del Conocimiento, Universidad Complutense de Madrid, 28040-Madrid, España.*

To cite this article: Sierra García, Jesús Enrique, Santos Peñas, Matilde. 2025. Wind turbine pitch control based on reinforcement learning and PID. XX Simposio CEA de Control Inteligente, Huelva (Spain), 2025.

Resumen

En este trabajo se propone una nueva estrategia de control que combina un controlador clásico como es el PID con una técnica de la inteligencia artificial, el aprendizaje por refuerzo. Esta arquitectura de control se aplica a control del ángulo de palas (pitch) de turbinas eólicas. El controlador basado en aprendizaje por refuerzo está focalizado a la eficiencia del aerogenerador, y calcula las recompensas a las acciones de control considerando el error de potencia y su derivada. Estas recompensas se emplean para actualizar la política del controlador. El controlador, teniendo en cuenta la política aprendida y el estado del sistema, genera una consigna del ángulo de las palas. A su vez el PID también genera otra consigna de pitch de forma simultánea. Un selector se encarga de seleccionar una de las dos consignas para llevarla finalmente al actuador de pitch. Los resultados de simulación muestran cómo el controlador basado en aprendizaje por refuerzo combinado con el PID proporciona un menor error de potencia respecto a la nominal, especialmente en los primeros episodios, contribuyendo a acelerar el aprendizaje.

Palabras clave: Aprendizaje por refuerzo, controlador PID, control, ángulo de las palas, aerogenerador, energía eólica

Paper Title in English, Bold Style

Abstract

This paper proposes a new control strategy that combines a traditional PID controller with reinforcement learning, an artificial intelligence technique. This control architecture is applied to blade angle (pitch) control in wind turbines. The reinforcement learning-based controller focuses on wind turbine efficiency and calculates rewards for control actions by considering the power error and its derivative. These rewards are used to update the controller policy. Taking into account the learned policy and the system state, the controller generates a blade angle setpoint. The PID also generates another pitch setpoint simultaneously. A selector chooses one of the two setpoints to be finally transferred to the pitch actuator. Simulation results show how the reinforcement learning-based controller combined with the PID provides a lower power error relative to the nominal power, especially in the early episodes, helping to accelerate learning.

Keywords: Reinforcement learning, PID regulator, control, pitch angle, wind turbine, wind energy.

1. Introducción

Los aerogeneradores son dispositivos complejos que aprovechan la energía eólica para generar electricidad limpia y renovable. El proceso implica la conversión de la energía cinética del viento en energía mecánica a través del giro del rotor y las palas, que luego se transforma en energía eléctrica mediante un generador [Buestán-Andrade et al., 2023]. Para optimizar la producción de energía, se emplean diversas estrategias de control que ajustan las variables de funcionamiento de la turbina eólica, como son la velocidad del generador, el ángulo de las palas (*pitch*) y la orientación del aerogenerador (control de guiñada o *yaw*), en función de las condiciones del viento incidente [Muñoz-Palomeque et al., 2024].

El control de pitch, en particular, presenta desafíos significativos debido a la naturaleza no lineal y compleja del sistema, así como a la imprevisibilidad del viento. Estos retos se acentúan en los aerogeneradores flotantes (FOWT, Floating Offshore Wind Turbine), situados en alta mar, que están expuestos a condiciones ambientales más severas. Para abordar estas dificultades se están explorando enfoques de control avanzados, incluyendo aquellos basados en inteligencia artificial y aprendizaje automático [Buestán-Andrade et al., 2024]. Entre estas técnicas, el aprendizaje por refuerzo (RL) se perfila como una estrategia prometedora, dada su capacidad para aprender de la experiencia y manejar problemas de control no lineales [Sierra-García y Santos, 2024], lo que podría mejorar significativamente la eficiencia y

*Autor para correspondencia: jesierra@ubu.es
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

la durabilidad de los aerogeneradores en diversos entornos operativos.

Existen otros trabajos en los que se emplea aprendizaje por refuerzo para el control de aerogeneradores. Por nombrar algunos, en [Zhang et al., 2019] se ofrece un resumen de aprendizaje por refuerzo basado en Deep learning para aplicaciones en sistemas de potencia. Una revisión más reciente sobre la aplicación de esta técnica a las turbinas eólica se presenta en Narayanan et al. [Narayanan, 2024]. Más específicamente para el ángulo colectivo de las palas, Hassan et al. [Hassan, 2025] muestran las propuestas de técnicas de la inteligencia artificial para el sistema de control de pitch.

En [Fernandez-Gauna et al., 2017], se utiliza RL para controlar turbinas de velocidad variable. Los mismos autores en un trabajo posterior proponen la utilización de RL condicionado para reducir el problema del crecimiento del espacio de estados [Fernandez-Gauna et al., 2018]. En [Abouheaf et al., 2018], se utiliza un modelo RL basado en el enfoque actor-crítico adaptativo para operar un generador de turbina de doble inducción. También se ha empleado en el control de yaw, por ejemplo, Saénz-Aguirre et al. presentan una estrategia de control basada en RL para el control de guiñada de una turbina [Saénz-Aguirre et al., 2019].

En este trabajo se propone una novedosa estrategia de control que combina RL y control convencional PID. El PID genera las consignas de control en los primeros episodios del entrenamiento, en los que el controlador basado en aprendizaje por refuerzo todavía no ha aprendido a controlar el sistema. De esta forma se consigue que el controlador aprenda más rápidamente a reducir el error de potencia.

El resto del artículo se estructura de la siguiente forma. La sección 2 describe las ecuaciones matemáticas del modelo de turbina utilizadas en la simulación. La estructura del controlador se detalla en la sección 3. La sección 4 explica los resultados de la aplicación del controlador a un caso de simulación. Finalmente, el artículo acaba con las conclusiones y los trabajos futuros.

2. Modelado de turbina

Para validar el funcionamiento del controlador se ha empleado un modelo electromecánico de una turbina pequeña de 7kW. El modelo recoge la transformación de la velocidad del viento en potencia. Esta potencia se transforma en par que acelera el rotor teniendo en cuenta la inercia *J*.

Por un lado, la ecuación del generador proporciona la variación de la corriente a partir de la velocidad del rotor (1). Como la turbina eólica es pequeña, no se emplea reductora. El aerogenerador está equipado con un actuador de pitch con constante de tiempo T_θ, modelado por (6). La forma aerodinámica de la pala determina el coeficiente de potencia Cp de la turbina. No es fácil desarrollar una expresión matemática del Cp por lo que normalmente se obtiene de forma empírica a partir de resultados experimentales ya que es específico para cada turbina. Estos datos experimentales pueden aproximarse mediante curvas de naturaleza exponencial como la de la ecuación (4).

Las ecuaciones que describen el comportamiento se presentan a continuación (1-7). Su desarrollo puede consultarse en [Sierra-García & Santos, 2020].

$$\dot{I}_a = \frac{1}{L_a}\left(K_g \cdot K_\phi \cdot w - (R_a + R_L) I_a\right) \tag{1}$$

$$\lambda = w \cdot \frac{R}{v} \tag{2}$$

$$\lambda_i = \left[\left(\frac{1}{\lambda + c_8}\right) - \left(\frac{c_9}{\theta^3 + 1}\right)\right]^{-1} \tag{3}$$

$$C_p(\lambda_i, \theta) = c_1 \left[\frac{C_2}{\lambda_i} - c_3\theta - c_4\theta^{c_5} - c_6\right] e^{-\frac{c_7}{\lambda_i}} \tag{4}$$

$$\dot{w} = \frac{1}{2 \cdot J \cdot w}\left(C_p(\lambda_i, \theta) \cdot \rho\pi R^2 \cdot v^3\right) - \frac{1}{J}\left(K_g \cdot K_\phi \cdot I_a + K_f w\right) \tag{5}$$

$$\ddot{\theta} = \frac{1}{T_\theta}\left[K_\theta\left(\theta_{ref} - \theta\right) - \dot{\theta}\right] \tag{6}$$

$$P_{out} = R_L \cdot {I_a}^2 \tag{7}$$

Donde L_a es la inductancia de la armadura del generador (H), K_g es una constante sin dimensiones del generador, K_ϕ es la constante de acoplamiento de flujo magnético (V·s/rad), R_a es la Resistencia de la armadura (Ω), R_L es la resistencia de la carga (Ω), considerada en el estudio como puramente resistiva, w es la velocidad angular del rotor (rad/s), I_a es la corriente en la armadura (A), λ es la relación de velocidad de punta, los valores de los coeficientes c_1 a c_9 dependen de las características de la turbina, J es la inercia rotacional (Kg.m2), R es el radio o longitud de la pala (m), ρ es la densidad del aire (Kg/m3), K_f es el coeficiente de fricción (N m/rad/s), θ_{ref} es la referencia para el actuador del pitch, θ es el pitch actual (rad), K_θ y T_θ son parámetros del actuador de pitch, v es la velocidad del viento (m/s).

La Tabla 1 muestra los parámetros empleados en las simulaciones [Mikati, et al., 2013].

Tabla 1: Métricas

Parámetro	Descripción	Valor
L_a	Inductancia de la armadura	13.5 mH
K_g	Constante del generador	23.31
K_ϕ	Constante de acoplamiento del flujo magnético	0.264 V/rad/s
R_a	Resistencia de la armadura	0.275 Ω
R_L	Resistencia de la carga	8 Ω
J	Inercia	6.53 Kg m^2
R	Radio de la pala	3.2 m
ρ	Densidad del aire	1.223 Kg/m^3
K_f	Coeficiente de fricción	0.025 N m/rad/s
$[c_1, c_2, c_3]$	C_p constantes	$[0.73, 151, 0.58]$
$[c_4, c_5, c_6]$	C_p constantes	$[0.002, 2.14, 13.2]$
$[c_7, c_8, c_9]$	C_p constantes	$[18.4, -0.02, -0.003]$
$[K_\theta, T_\theta]$	Constantes del actuador de pitch	[0.15, 2]

3. Arquitectura de control

La arquitectura de control propuesta se muestra en la Figura 1. Esta constituida por dos controladores: un regulador PID y un controlador basado en aprendizaje por refuerzo (RL). El PID genera la referencia para el actuador de pitch θ_{PID}, mientras que el controlador basado en aprendizaje por refuerzo genera la referencia denotada como θ_{RL}. Estas dos referencias no se llevan simultáneamente al aerogenerador, sino que se emplea un selector que decide en qué momento se aplica una u otra.

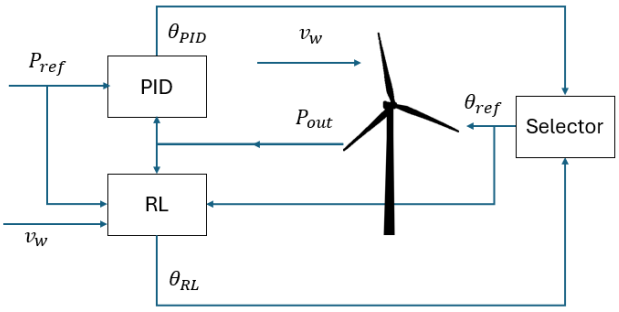


Figura 1: Arquitectura de control.

Las referencias de pitch producen un cambio en el sistema que hacen que el error de potencia P_{err} aumente o disminuya. Siguiendo la nomenclatura propia del aprendizaje por refuerzo [Vamvoudakis, 2021], en base a ese error el controlador RL calcula las recompensas r_t de las acciones a_{t-1}. Estas recompensas se asocian a las acciones llevadas a cabo por sí mismo o por el controlador PID. Para ello recibe como entrada la señal θ_{ref} del módulo Selector que recoge ambas acciones. Estas recompensas se utilizan para actualizar la política de control. De esta forma el controlador aprende de las acciones tomadas por el controlador PID y de las acciones exploradas por sí mismo.

El controlador RL, teniendo en cuenta el estado del sistema, que se ha definido por las tres variables: viento, error de potencia, y derivada de error de potencia, y junto con la política aprendida, genera las consignas de pitch θ_{RL}

El controlador puede formalizarse con las siguientes ecuaciones:

$$P_{err}(t_i) = P_{ref}(t_{i-1}) - P_{out}(t_{i-1}) \tag{8}$$

$$s_t(t_i) = f_d[P_{err}(t_i), \dot{P}_{err}(t_i), v_w(t_i)] \tag{9}$$

$$r(t_i) = \begin{cases} -dP_{\mathrm{err_S}}(t_i)/dP_{\mathrm{err_{MAX}}} & P_{\mathrm{err}}(t_{i-1}) > 0 \\ dP_{\mathrm{err_S}}(t_i)/dP_{\mathrm{err_{MAX}}} & P_{\mathrm{err}}(t_{i-1}) < 0 \end{cases} \tag{10}$$

$$Q_{(s_{t-1},a_{t-1})}(t_i) = f_\Pi(Q_{(s_{t-1},a_{t-1})}(t_i), r(t_i)) \tag{11}$$

$$a(t_i) = arg\max_a(Q_{(s_t,a)}) \tag{12}$$

$$\theta_{RL}(t_i) = a(t_i) \cdot \frac{\pi/2}{n_{act}} \tag{13}$$

$$\theta_{PID}(t_i) = \frac{\pi}{2} - G[K_p P_{err} + K_p \dot{P}_{err} + K_I \int P_{err}] \tag{14}$$

$$\theta_{ref}(t_i) = \begin{cases} \theta_{RL}(t_i), & sel = 0 \\ \theta_{PID}(t_i), & sel = 1 \end{cases} \tag{15}$$

$$a_{t-1} = \theta_{ref}(t_i)\frac{n_{act}}{\pi/2} \tag{16}$$

$$s_{t-1} = s_t(t_i) \tag{17}$$

En las ecuaciones anteriores, s_t denota el estado en el tiempo t_i , f_d es una función de discretización de estados, f_Π denota la función de actualización de la política; por ejemplo, puede usarse Q-learning para implementar esta función. La ecuación (16) se emplea para que el controlador aprenda de la última acción ejecutada, independientemente de si ha sido tomada por el PID o por el RL.

4. Resultados

El controlador se ha validado en simulación empleando el software Matlab con un perfil de viento mostrado en la Figura 2. Como se puede apreciar, se trata de un perfil de viento en diente de sierra con valores de viento comprendido entre 12.25 m/s y 13 m/s, que está dentro del rango de valores de viento que corresponden a la región de operación nominal de la turbina, donde es necesario aplicar el control del ángulo de palas para mantener la velocidad de rotación de las palas constante y por lo tanto producir una potencia constante en su valor máximo.

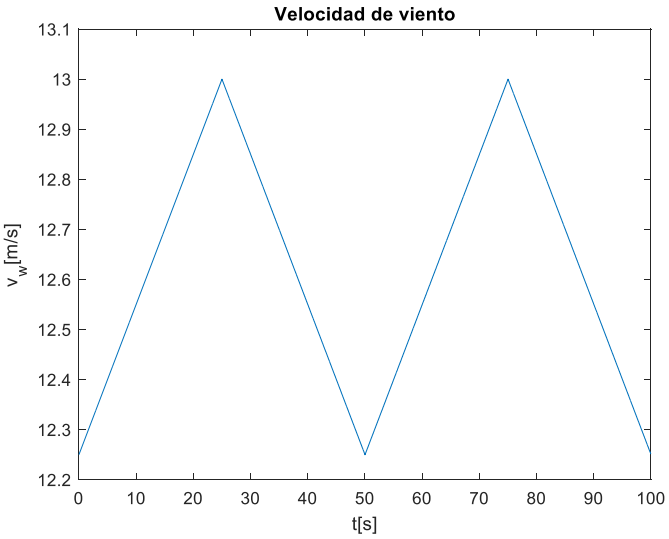


Figura 2: Perfil de viento.

El controlador se ha comparado con un controlador basado en aprendizaje por refuerzo y con un controlador PID. El controlador RL puro se obtiene manteniendo el valor de la señal del selector $sel = 0$ de forma permanente. El controlador PID se obtiene manteniendo el valor de la señal $sel = 1$ de forma permanente (ecuación 15).

En este trabajo se ha empleado la siguiente configuración para ambos controladroes: $n_{act} = 100$, G=7.85e-4, para el

control con aprendizaje reforzado, y parámetros del PID Kp=1, Kd=0.2, Ki=0.1. La señal sel se ha fijado a 1 durante el primer episodio (actúa el PID) y a 0 (actúa el controlador RL) durante el resto de los episodios.

La Figura 3 muestra la comparación de la evolución del error RMSE de potencia para los diferentes controladores evaluados. Se observa como el RMSE se reduce mientras los controladores basados en aprendizaje por refuerzo están aprendiendo. Ambos controladores, RL (línea azul) y RL-PID (línea roja), consiguen alcanzar valores inferiores de error al PID (línea amarilla) tras 6 episodios, por lo que producen una potencia más cercana al máximo una vez han aprendido.

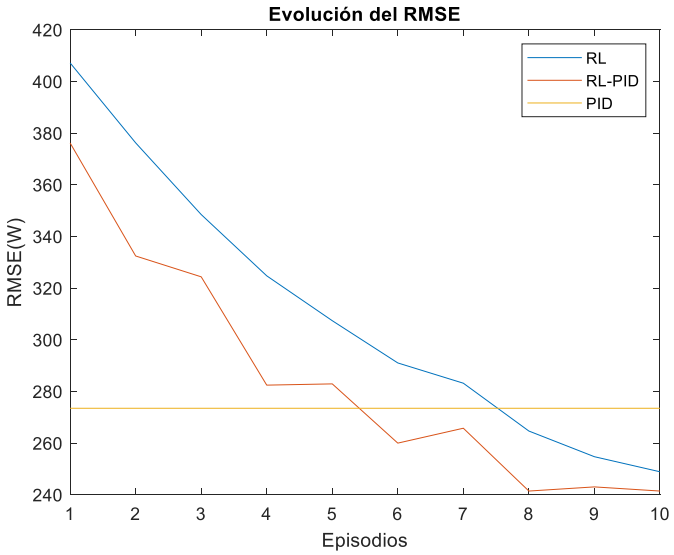


Figura 3: Evolución del RMSE de la potencia .

Es interesante recalcar el hecho de que el controlador que combina RL y PID consigue reducir el error durante los primeros episodios gracias al aprendizaje con las consignas generadas por el PID.

La Figura 4 muestra la comparación de las señales de potencia de salida obtenidas con los diferentes controladores. La señal morada representa la potencia que se obtiene cuando el ángulo de pitch se fija de forma permanente a cero grados, y la verde cuando se fija de forma permanente a pi/2 grados. Estas señales de potencia conforman el margen de funcionamiento del sistema, puesto que no se puede conseguir una potencia mayor a la señal morada ni una menor que la señal verde.

Es posible observar como el peor rendimiento es el obtenido con el controlador PID (línea amarilla), ya que la potencia se aleja más de la referencia, sobre todo en el intervalo de 30 a 50 segundos. En algunos puntos el controlador basado en aprendizaje por refuerzo puro (línea azul) funciona peor que el PID (intervalo de 50 a 75 segundos). Sin embargo, en este caso, el RL-PID (línea roja) consigue imitar razonablemente bien el comportamiento del PID y, en consecuencia, proporciona un menor error en su conjunto, siendo el controlador que proporciona un mejor comportamiento, es decir, que está más cerca de la potencia nominal (línea negra discontinua) la mayor parte del tiempo.

Un aspecto importante de analizar es la acción de control. La Figura 5 muestra la comparación de la señal de pitch obtenida con los diferentes controladores.

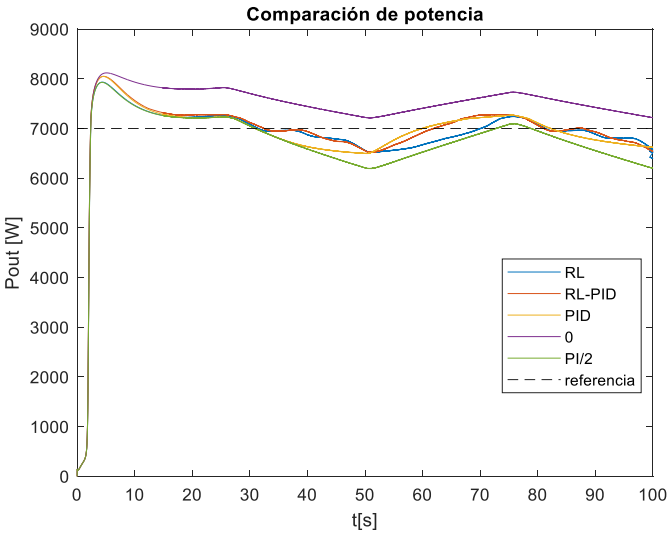


Figura 4: Comparación de las señales de potencia obtenidas con los diferentes controladores.

En la Figura 5 se puede observar un resultado interesante ya que las señales RL (línea azul) y RL-PID (línea roja) se parecen mucho en el intervalo 30-50 segundos, que coincide con los valores en los que ambos controladores que usan el aprendizaje por refuerzo proporcionan mejores resultados que el PID (línea amarilla). Sin embargo, en el intervalo 50-75 segundas la señal RL-PI se aleja de la señal del control RL y se aproxima mas al PID, que es el que funciona mejor en ese tramo.

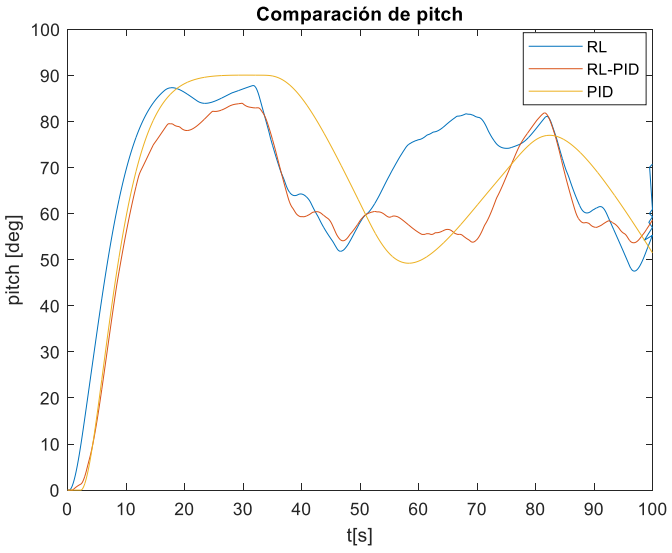


Figura 5: Comparación de las señales de pitch obtenidas con los diferentes controladores.

5. Conclusiones y trabajos futuros

En ente trabajo se presenta una nueva estrategia de control que combina el control basado en PID con el aprendizaje por refuerzo. Este enfoque híbrido combina las ventajas del control PID tradicional con la capacidad de adaptación y optimización del aprendizaje por refuerzo. El controlador PID proporciona una base estable y confiable durante las etapas iniciales del proceso de aprendizaje, garantizando un rendimiento aceptable

mientras el agente de RL acumula experiencia. A medida que el controlador basado en RL mejora su desempeño a través de la interacción con el sistema, gradualmente asume un papel más prominente en la generación de señales de control.

La integración de estas dos técnicas ofrece varios beneficios. Primero, reduce el tiempo de aprendizaje necesario para que el controlador RL alcance un rendimiento óptimo, ya que comienza con una política de control razonable proporcionada por el PID. Segundo, mitiga los riesgos asociados con las acciones exploratorias potencialmente peligrosas del agente RL durante las etapas iniciales de aprendizaje. Además, esta estrategia permite una transición suave del control convencional al control basado en técnicas de la inteligencia artificial, lo que puede facilitar la adopción de técnicas de aprendizaje por refuerzo en aplicaciones industriales donde la confiabilidad y la seguridad son primordiales.

Como principal línea futura cabe destacar el diseño de nuevas estrategias para la combinación del aprendizaje por refuerzo con controladores clásicos. En la medida de lo posible también se validará la propuesta de control con un prototipo a escala de una turbina eólica.

Agradecimientos

Este trabajo fue parcialmente financiado por el Ministerio de Ciencia e Innovación de España, como parte del proyecto MCI/AEI/FEDER número PID2021-123543OB-C21, y la Ayuda RED2022-134588-T financiada por MICIU/AEI/10.13039/501100011033.

Referencias

Abouheaf, M., Gueaieb, W., & Sharaf, A. (2018). Model-free adaptive learning control scheme for wind turbines with doubly fed induction generators. IET Renewable Power Generation, 12(14), 1675-1686.

Buestán-Andrade, P. A., Santos, M., Sierra-García, J. E., & Pazmiño-Piedra, J. P. (2023, August). Comparison of LSTM, GRU and transformer neural network architecture for prediction of wind turbine variables. In International conference on soft computing models in industrial and environmental applications (pp. 334-343). Cham: Springer Nature Switzerland.

Buestán-Andrade, P. A., Peñacoba-Yagüe, M., Sierra-García, J. E., & Santos, M. (2024). Wind power forecasting with machine learning algorithms in low-cost devices. Electronics, 13(8), 1541.

Fernandez-Gauna, B., Fernandez-Gamiz, U., & Grana, M. (2017). Variable speed wind turbine controller adaptation by reinforcement learning. Integrated Computer-Aided Engineering, 24(1), 27-39

Fernandez-Gauna, B., Osa, J. L., & Graña, M. (2018). Experiments of conditioned reinforcement learning in continuous space control tasks. Neurocomputing, 271, 38-47

Hassan, A., Ahmad, G., Shafiullah, M., Islam, A., & Alam, M. S. (2025). Review of the Intelligent Frameworks for Pitch Angle Control in Wind Turbines. IEEE Access.

Mikati, M., Santos, M., Armenta, C. (2013). Electric grid dependence on the configura-tion of a small-scale wind and solar power hybrid system. Renewable energy 57, 587–593

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.

Narayanan, V. L. (2024). Reinforcement learning in wind energy-a review. International Journal of Green Energy, 21(9), 1945-1968.

Saénz-Aguirre, A., Zulueta, E., Fernández-Gamiz, U., Lozano, J., Lopez-Guede, J. M. (2019). Artificial neural network based reinforcement learning for wind turbine yaw control. Energies,12(3), 436.

Sierra-García, J. E., & Santos, M. (2020). Performance analysis of a wind turbine pitch neurocontroller with unsupervised learning. Complexity, 2020

Sierra-Garcia, J. E., & Santos, M. (2024). Federated discrete reinforcement learning for automatic guided vehicle control. Future Generation Computer Systems, 150, 78-89.

Vamvoudakis, K. G., Wan, Y., Lewis, F. L., & Cansever, D. (Eds.). (2021). Handbook of reinforcement learning and control. Springer International Publishing.

Zhang, Z., Zhang, D., & Qiu, R. C. (2019). Deep reinforcement learning for power sys-tem applications: An overview. CSEE Journal of Power and Energy Systems, 6(1), 213-225.