

Optimal control for wind turbine based on reinforcement learning

Siame Chouiekh^{1,}, Yassamine Zoubaa¹, Ayoub El Bakri¹, and Ismail Boumhidi¹*

¹LISAC Laboratory, Dept. of Physics, Faculty of Sciences Sidi Mohamed Ben Abdellah University Fez, Morocco.

Abstract. In this conference paper, an optimal control method is designed for a variable speed wind turbine system. Due to the inherent nonlinearity of the wind turbine arising from the aerodynamic torque, a linearized model is derived to handle the system's nonlinearities. An online update cost function is created based on the resulting linearized model. The critic neural network weight vector is updated with the steepest decent algorithm to design an optimal control able to minimize the given cost function. To validate the effectiveness of the optimal control based on reinforcement learning, simulation results with varying wind speed profile for different values of learning parameters are presented.

1 Introduction

Several research teams are currently interested in sustainable development and renewable energies. Thus, the wind turbine development represents a significant investment in technology research. The systems that generate electrical energy from the wind can provide a technological and economic alternative to the various sources of exhaustible energy. Therefore, the study of wind turbine systems is more important and necessary. In general, wind turbine systems fall into two categories: fixed-speed wind turbine systems and variable-speed wind turbine systems. Due to its high efficiency and the optimization of electricity production in the grid at different wind speeds, the variable-speed wind turbines have become very popular compared to fixed speed wind turbines [1, 2]. Strong nonlinear characteristics of the variable speed wind turbine, as is well known, can make the wind turbine's operation more challenging. Thus, the linearization around an operating point of this system has been made in [3]–[5].

Optimal control, as is well known, leads controlled plants to a specific optimality criterion, which is always important in contemporary complex systems. In other words, it is focused on identifying a control strategy that optimally guides the dynamical system to the equilibrium point, with regard to a cost function. Hamilton-Jacobi-Bellman (HJB) equations are solved for nonlinear system to obtain optimal control.

The use of adaptive dynamic programming (ADP) is a powerful and significant method to solve HJB equations [6]–[10]. Therefore, it is generally used to find the optimal control law in recent years. Reinforcement learning (RL), neuron-dynamic programming, adaptive

* Corresponding author: siame.chouiekh@usmba.ac.ma

critic designs and neural dynamic programming are all synonyms for ADP. The critic neural network is one of intelligent technologies of ADP, has been used in [11]–[15] to estimate the iterative cost function (performance index function) and control law. In this context, this paper, investigate the performances of the integration of RL in the control of wind turbine systems.

The structure of this conference paper consists of the following sections. In section 2, modeling and linearization of variable speed wind turbine is discussed. In section 3, the optimal control applied to the linearized model of a variable speed wind turbine is shown. In order to validate the proposed optimal control method, the simulations of rotor speed for different learning parameters are shown in section 4. This paper is ended with a conclusion.

2 Modelling and linearization of wind turbine system

2.1 Modelling of wind turbine

A wind turbine system consists of different interconnecting subsystems: the mechanical subsystem, generator subsystem, and converter subsystem. In this study, we focus on the mechanical subsystem, which includes the aerodynamic subsystem and the transmission subsystem (drive train subsystem).

The aerodynamic subsystem shows how wind speed is converted as forces acting on the blades that generate rotational motion. The power captured by the rotor is given by:

$$P_r = 0.5\rho\pi R^2 v^3 C_p(\lambda, \beta) \quad (1)$$

Where ρ , v and R are the air density, the wind speed m/s, and the rotor radius, respectively. And C_p is the power coefficient, which is dependent on both the pitch angle β and the tip speed ratio λ is given as follows:

$$\lambda = \frac{\omega_r R}{v} \quad (2)$$

With ω_r is rotor speed (rad/s). The aerodynamic torque expression is:

$$T_{ae} = 0.5\rho\pi R^3 v^2 C_q(\lambda, \beta) \quad (3)$$

$$C_q(\lambda, \beta) = \frac{C_p(\lambda, \beta)}{\lambda} \quad (4)$$

Where $C_q(\lambda, \beta)$ is the torque coefficient.

The C_q can be approximated using a modified and expanded form of the function for the power coefficient given in [16].

$$C_q(\lambda, \beta) = c_1(1 + c_2(\beta + c_3)^{0.5}) + \frac{c_4}{\lambda}(c_5\lambda_l(\lambda, \beta) - c_6\beta - c_7\beta^{c_8} - c_9)\exp(-c_{10}\lambda_l(\lambda, \beta)) \quad (5)$$

$$\text{With } \lambda_l(\lambda, \beta) = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{c_{11} + c_{12}\beta^3} \quad (\lambda > 0)$$

The transmission subsystem transfers the mechanical power taken by the rotor to the electric machine. The following dynamic system represent the transmission between the rotor and the generator:

$$\begin{bmatrix} \dot{\omega}_r \\ \dot{\omega}_g \\ \dot{T}_{ls} \end{bmatrix} = \begin{bmatrix} -\frac{K_r}{J_r} & 0 & -\frac{1}{J_r} \\ 0 & -\frac{K_g}{J_g} & \frac{1}{n_g J_g} \\ B_{ls} - \frac{K_{ls} K_r}{J_r} & \frac{1}{n_g} \left(\frac{K_{ls}}{J_g} - B_{ls} \right) & -K_{ls} \left(\frac{J_r + n_g^2 J_g}{n_g^2 J_g J_r} \right) \end{bmatrix} \begin{bmatrix} \omega_r \\ \omega_g \\ T_{ls} \end{bmatrix} + \begin{bmatrix} \frac{1}{J_r} \\ 0 \\ \frac{K_{ls}}{J_r} \end{bmatrix} T_{ae} + \begin{bmatrix} 0 \\ -\frac{1}{J_g} \\ \frac{K_{ls}}{n_g J_g} \end{bmatrix} T_{em} \quad (6)$$

2.2 Linearization of wind turbine system

The nonlinearity of wind turbine model (6) comes from aerodynamic torque T_{ae} , which depends on the wind speed $v(t)$, the pitch angle β and rotor speed ω_r . The linearization of aerodynamic torque T_{ae} around a given operating point is given by the following expression:

$$T_{ae}(\omega_r, \beta, v) = T_{ae}(\omega_{r_{op}}, \beta_{op}, v_{op}) + \Delta T_{ae} \quad (7)$$

$$\Delta T_{ae} = \alpha \Delta \omega_r + \gamma \Delta \beta + \delta \Delta v \quad (8)$$

With

$$\alpha = \left. \frac{\partial T_{ae}}{\partial \omega_r} \right|_{op} = 0.5 \rho \pi R^4 v_{op} \left. \frac{\partial C_q(\lambda, \beta)}{\partial \lambda} \right|_{op} ; \quad \gamma = \left. \frac{\partial T_{ae}}{\partial \beta} \right|_{op} = 0.5 \rho \pi R^3 v_{op}^2 \left. \frac{\partial C_q(\lambda, \beta)}{\partial \beta} \right|_{op}$$

$$\delta = \left. \frac{\partial T_{ae}}{\partial v} \right|_{op} = 0.5 \rho \pi R^3 v_{op} \left[2 C_q(\lambda_{op}, \beta_{op}) - \lambda_{op} \left. \frac{\partial C_q(\lambda, \beta)}{\partial \lambda} \right|_{op} \right]$$

The variations around the operating point are defined as follows:

$$\Delta \omega_r = \omega_r - \omega_{r_{op}} ; \quad \Delta \omega_g = \omega_g - \omega_{g_{op}} ; \quad \Delta v = v - v_{op} ;$$

$$\Delta \beta = \beta - \beta_{op} ; \quad \Delta T_{em} = T_{em} - T_{em_{op}} ; \quad \Delta T_{ls} = T_{ls} - T_{ls_{op}}$$

By taking $\Delta \omega_r$, $\Delta \omega_g$, and ΔT_{ls} are the new state variables, and using system (6) and equation (8), the linearized state space model is given by:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A} \mathbf{x}(t) + \mathbf{B} \mathbf{u}(t) + \mathbf{D} \Delta v \\ \mathbf{y} = \mathbf{C} \mathbf{x}(t) \end{cases} \quad (9)$$

With

$\mathbf{x} = [\Delta \omega_r, \Delta \omega_g, \Delta T_{ls}]^T$, $\mathbf{y} = \Delta \omega_r$, $\mathbf{u}(t) = [\Delta T_{em}, \Delta \beta]^T$ is the control input.

$$\mathbf{A} = \begin{bmatrix} \frac{\alpha - K_r}{J_r} & 0 & -\frac{1}{J_r} \\ 0 & -\frac{K_g}{J_g} & \frac{1}{n_g J_g} \\ (B_{ls} - \frac{K_{ls} K_r}{J_r}) + \frac{\alpha K_{ls}}{J_r} & \frac{1}{n_g} \left(\frac{K_{ls}}{J_g} - B_{ls} \right) & -K_{ls} \left(\frac{J_r + n_g^2 J_g}{n_g^2 J_g J_r} \right) \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & \frac{\gamma}{J_r} \\ -\frac{1}{J_g} & 0 \\ \frac{K_{ls}}{n_g J_g} & \frac{\gamma K_{ls}}{J_r} \end{bmatrix},$$

$$\mathbf{D} = \begin{bmatrix} \frac{\delta}{J_r} \\ 0 \\ \frac{\delta K_{ls}}{J_r} \end{bmatrix}, \quad \mathbf{C} = [0 \ 1 \ 0]$$

3 Controller design of wind turbine

For simplicity in designing a control law based on reinforcement learning, an affine form of system (9) is written as:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) \mathbf{u}(t) \quad (10)$$

With $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}(t) + \mathbf{D}\Delta v$ and $\mathbf{g}(\mathbf{x}) = \mathbf{B}$.

In this section, an optimal control law $\mathbf{u}^*(t)$ for system (10) is designed to minimize the cost function below:

$$\mathcal{V}(\mathbf{x}) = \int_0^\infty (\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u}) dt \quad (11)$$

$\mathbf{R}$ and $\mathbf{Q}$ are positive definite matrixes.

If $\mathcal{V}(\mathbf{x})$ is continuously differentiable, the following expression describes its infinitesimal version:

$$\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u} + \nabla_{\mathbf{x}} \mathcal{V}^T (\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) \mathbf{u}) = 0 \quad (12)$$

With $\mathcal{V}(0) = 0$, and $\nabla_{\mathbf{x}} \mathcal{V} = \frac{\partial \mathcal{V}(\mathbf{x})}{\partial \mathbf{x}}$. Define the Hamiltonian function of optimal cost function and the optimal control problem as:

$$H(\nabla_{\mathbf{x}} \mathcal{V}, \mathbf{x}, \mathbf{u},) = \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u} + \nabla_{\mathbf{x}} \mathcal{V}^T (\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) \mathbf{u}) \quad (13)$$

When the cost function $\mathcal{V}(\mathbf{x})$ converges to its optimal value $\mathcal{V}^*(\mathbf{x})$, the following HJB equation is satisfied.

$$\min (H(\nabla_{\mathbf{x}} \mathcal{V}^*, \mathbf{x}, \mathbf{u})) = 0 \quad (14)$$

If the solution $\mathcal{V}^*(\mathbf{x})$ exists and is continuously differentiable, the optimal control $\mathbf{u}^*$ can be described as:

$$\mathbf{u}^*(t) = -0.5 \mathbf{R}^{-1} \mathbf{g}^T(\mathbf{x}) \nabla_{\mathbf{x}} \mathcal{V}^*(\mathbf{x}, t) \quad (15)$$

It is hard to find the solution of equation (14) through dynamic programming approaches. Thus, a critic neural network is employed to estimate the cost function $\mathcal{V}(\mathbf{x})$ as:

$$\mathcal{V}(\mathbf{x}) = \mathbf{W}_c^T \boldsymbol{\sigma}(\mathbf{x}) + E(\mathbf{x}) \quad (16)$$

Where $\mathbf{W}_c \in \mathbb{R}^n$ is the weight vector of the critic neural network, $\boldsymbol{\sigma}(\mathbf{x}) \in \mathbb{R}^n$ is the activation function, $E(\mathbf{x})$ is the neural network error, and n is the number of neurons in the hidden layer. The gradient of equation (16) with respect to the system state $\mathbf{x}$ is:

$$\nabla_{\mathbf{x}} \mathcal{V}(\mathbf{x}) = (\nabla \boldsymbol{\sigma}(\mathbf{x}))^T \mathbf{W}_c + \nabla E(\mathbf{x}) \quad (17)$$

With $\nabla \sigma(\mathbf{x}) = \frac{\partial \sigma(\mathbf{x})}{\partial \mathbf{x}}$, and $\nabla E(\mathbf{x})$ are the gradients of the activation function and the neural network error respectively.

Therefore, the Hamiltonian (13) becomes as:

$$H(\mathbf{x}, \mathbf{u}, \mathbf{W}_c) = \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u} + \mathbf{W}_c^T \nabla \sigma(\mathbf{x}) \dot{\mathbf{x}} + \nabla E(\mathbf{x}) \dot{\mathbf{x}} = \varepsilon_c \quad (18)$$

The critic neural network (16) can be approximated as follows:

$$\hat{V}(\mathbf{x}) = \hat{\mathbf{W}}_c^T \sigma(\mathbf{x}) \quad (19)$$

Where $\hat{\mathbf{W}}_c$ is the estimated weight. The gradient of (19) with respect to the system state $\mathbf{x}$ is:

$$\nabla_{\mathbf{x}} \hat{V}(\mathbf{x}) = (\nabla \sigma(\mathbf{x}))^T \hat{\mathbf{W}}_c \quad (20)$$

Thus, the estimate Hamiltonian can be expressed as:

$$H(\mathbf{x}, \mathbf{u}, \hat{\mathbf{W}}_c) = \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u} + \hat{\mathbf{W}}_c^T \nabla \sigma(\mathbf{x}) \dot{\mathbf{x}} = e_c \quad (21)$$

Therefore, the objective function $\xi_c = \frac{1}{2} e_c^T e_c$ is minimized by using the steepest decent algorithm, the critic neural network weight vector $\hat{\mathbf{W}}_c$ can be updated as:

$$\dot{\hat{\mathbf{W}}}_c = -l_c \left(e_c - \mu \operatorname{sign}(\hat{\mathbf{W}}_c^T \boldsymbol{\theta}) \right) \boldsymbol{\theta} \quad (22)$$

Where $\boldsymbol{\theta} = \nabla \sigma(\mathbf{x}) \dot{\mathbf{x}}$

Where $l_c > 0$ is the learning rate of critic neural network (22) and $\mu > 0$ is a designed parameter. Therefore, the optimal control law define as:

$$\hat{\mathbf{u}}(t) = -0.5 \mathbf{R}^{-1} \mathbf{g}^T(\mathbf{x}) (\nabla \sigma(\mathbf{x}))^T \hat{\mathbf{W}}_c \quad (23)$$

$\hat{\mathbf{u}}$ converges to $\mathbf{u}^*$ when the neural network weights converge to ideal weights.

4 Simulation results

The numerical simulations is provided in this section to explain the effectiveness of the proposed optimal control, which is implemented in MATLAB SIMULINK. The characteristics of studied wind turbine are presented below:

$$\begin{aligned} R &= 63 \text{m}; \rho = 1.225 \text{Kg.m}^{-3}; K_r = 27.36 \text{ N} \cdot \frac{\text{m}}{\text{rad}} / \text{s}; J_r = 3.25 \times 10^5 \text{Kg.m}^2; \\ K_g &= 0.2 \text{ N} \cdot \frac{\text{m}}{\text{rad}} / \text{s}; J_g = 34.4 \text{Kg.m}^2; n_g = 97; K_{ls} = 9500 \text{ N} \cdot \frac{\text{m}}{\text{rad}} / \text{s}; \\ B_{ls} &= 2.691 \times 10^5 \text{N} \cdot \frac{\text{m}}{\text{rad}} / \text{s}. \end{aligned}$$

The wind speed profile used is shown in the Fig 1.

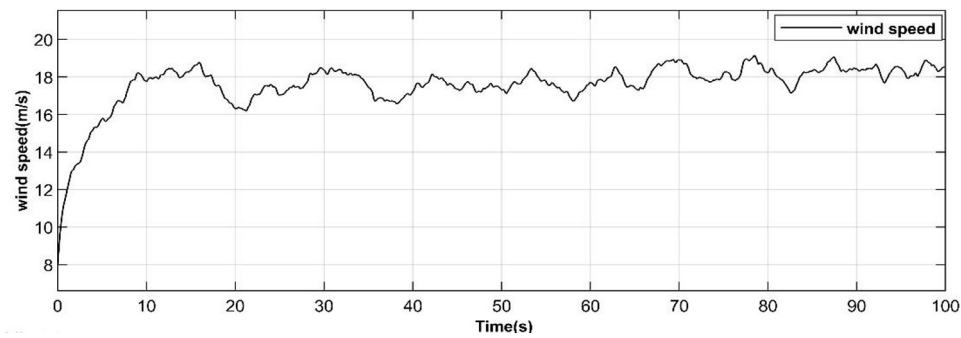


Fig. 1. Wind speed profile with a mean of 18m/s.

Let $\mathbf{Q} = 100 * \text{eye}(3)$ and $\mathbf{R} = [0.001 \ 0; 0 \ 100]$. In this simulation, a critic neural network has three input neurons, six hidden neurons, and one output neuron, it is used to estimate the cost function. The activation function is selected as $\sigma(\mathbf{x}) = [x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2]$. The figures 2-4 show the rotor speed of wind turbine and that desired for different values of the learning parameters (l_c and μ).

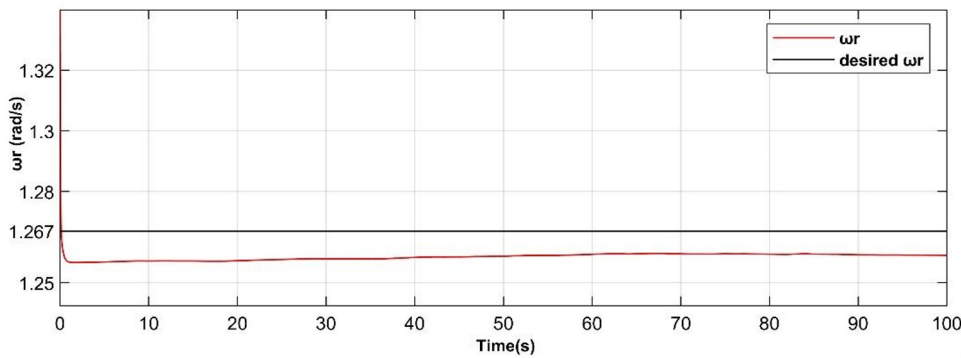


Fig. 2. Rotor speed for $l_c = 0.0001$ and $\mu = 1$.

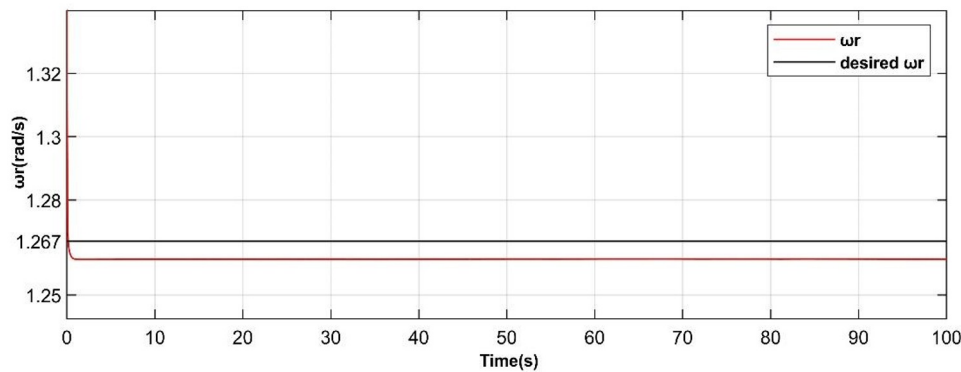


Fig. 3. Rotor speed for $l_c = 0.001$ and $\mu = 0.1$.

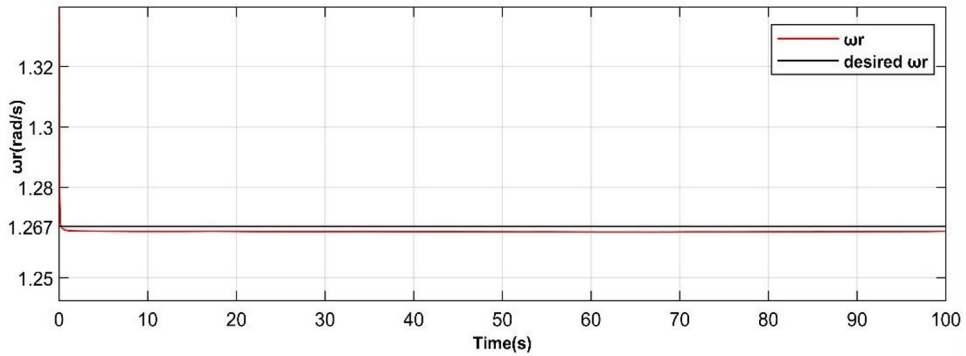


Fig. 4. Rotor speed for $l_c = 0.01$ and $\mu = 0.5$.

One may conclude from these plots that the proposed controller converges to the equilibrium point with an acceptable tracking error. After conducting several simulations, the optimal parameters were determined to be $l_c = 0.01$ and $\mu = 0.5$. Figure 3 illustrates that the rotor speed converges precisely to its desired value. This outcome serves as a validation of the effectiveness of the proposed optimal control approach.

5 Conclusion

In this paper, an optimal control method for a wind turbine system was studied. First, a linearization around an operating point of aerodynamic torque was made. Then, the adaptive dynamic programming methods are used to solving the optimal control problem. The online cost function was constructed, and subsequently a critic neural network was used to obtain the optimal cost function and the optimal control. Finally, the simulation results for different value of learning parameters were represented to test the efficacy of the proposed optimal control method. The simulation results demonstrate the performance of the proposed optimal control for the wind turbine system, and the significance of selecting the appropriate parameters. The main focus of this paper is on developing reinforcement learning based optimal control for the wind turbine, and the proposed method is evaluated based on the simulation results.

References

- [1] P. Anbalagan et Y. H. Joo, « Dissipative-based Sampled-data Control for T-S Fuzzy Wind Turbine System Via Fragmented-delayed State Looped Functional Approach », In Review, preprint, avr. 2022. doi: 10.21203/rs.3.rs-1525158/v1.
- [2] S. Georg et H. Schulte, « Takagi-Sugeno Sliding Mode Observer with a Weighted Switching Action and Application to Fault Diagnosis for Wind Turbines », in *Intelligent Systems in Technical and Medical Diagnostics*, vol. 230, J. Korbicz et M. Kowal, Éd., in *Advances in Intelligent Systems and Computing*, vol. 230. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, p. 41-52. doi: 10.1007/978-3-642-39881-0_3.
- [3] B. Boukhezzar et H. Siguerdidjane, « Comparison between linear and nonlinear control strategies for variable speed wind turbines », *Control Eng. Pract.*, vol. 18, n° 12, p. 1357-1368, déc. 2010, doi: 10.1016/j.conengprac.2010.06.010.

- [4] A. EL Bakri et I. Boumhidi, « Fuzzy model-based faults diagnosis of the wind turbine benchmark », *Procedia Comput. Sci.*, vol. 127, p. 464-470, 2018, doi: 10.1016/j.procs.2018.01.144.
- [5] S. Bougdour, R. Elhouti, S. Sefriti, et I. Boumhidi, « Optimal predictive control model of wind turbine », in *2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS)*, Fez, Morocco: IEEE, oct. 2021, p. 1-7. doi: 10.1109/ICDS53782.2021.9626757.
- [6] D. Liu, Q. Wei, D. Wang, X. Yang, et H. Li, *Adaptive Dynamic Programming with Applications in Optimal Control*. in Advances in Industrial Control. Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-50815-3.
- [7] R. Tang, B. Luo, et Y. Liao, « Adaptive dynamic programming based composite control for profile tracking with multiple constraints », *Neurocomputing*, vol. 557, p. 126711, nov. 2023, doi: 10.1016/j.neucom.2023.126711.
- [8] L. Kong, S. Zhang, et X. Yu, « Approximate optimal control for an uncertain robot based on adaptive dynamic programming », *Neurocomputing*, vol. 423, p. 308-317, janv. 2021, doi: 10.1016/j.neucom.2020.10.012.
- [9] G. Hu, J. Guo, J. Cieslak, Y. Ding, Z. Guo, et D. Henry, « Fault-tolerant control based on adaptive dynamic programming for reentry vehicles subjected to state-dependent actuator fault », *Eng. Appl. Artif. Intell.*, vol. 123, p. 106450, août 2023, doi: 10.1016/j.engappai.2023.106450.
- [10] X. Li, L. Wang, Y. An, Q.-L. Huang, Y.-H. Cui, et H.-S. Hu, « Dynamic path planning of mobile robots using adaptive dynamic programming », *Expert Syst. Appl.*, vol. 235, p. 121112, janv. 2024, doi: 10.1016/j.eswa.2023.121112.
- [11] B. Zhao, G. Shi, et D. Wang, « Asymptotically stable critic designs for approximate optimal stabilization of nonlinear systems subject to mismatched external disturbances », *Neurocomputing*, vol. 396, p. 201-208, juill. 2020, doi: 10.1016/j.neucom.2018.08.092.
- [12] A. Ziaei, H. Kharrati, et A. Rahimi, « Fault-tolerant control for nonlinear offshore steel jacket platforms based on reinforcement learning », *Ocean Eng.*, vol. 246, p. 110247, févr. 2022, doi: 10.1016/j.oceaneng.2021.110247.
- [13] B. Zhao, D. Liu, et Y. Li, « Observer based adaptive dynamic programming for fault tolerant control of a class of nonlinear systems », *Inf. Sci.*, vol. 384, p. 21-33, avr. 2017, doi: 10.1016/j.ins.2016.12.016.
- [14] R. Song et F. L. Lewis, « Robust optimal control for a class of nonlinear systems with unknown disturbances based on disturbance observer and policy iteration », *Neurocomputing*, vol. 390, p. 185-195, mai 2020, doi: 10.1016/j.neucom.2020.01.082.
- [15] B. Zhao, D. Liu, X. Yang, et Y. Li, « Observer-critic structure-based adaptive dynamic programming for decentralised tracking control of unknown large-scale nonlinear systems », *Int. J. Syst. Sci.*, vol. 48, n° 9, p. 1978-1989, juill. 2017, doi: 10.1080/00207721.2017.1296982.
- [16] S. Georg, H. Schulte, et H. Aschemann, « Control-oriented modelling of wind turbines using a Takagi-Sugeno model structure », in *2012 IEEE International Conference on Fuzzy Systems*, Brisbane, Australia: IEEE, juin 2012, p. 1-8. doi: 10.1109/FUZZ-IEEE.2012.6251302.