



Data-Driven Control for Proton Exchange Membrane Fuel Cells: Method and Application

Jiawen Li^{1*}, Kedong Zhu² and Tao Yu¹

¹College of Electric Power, South China University of Technology, Guangzhou, China, ²China Electric Power Research Institute (Nanjing), Nanjing, China

A data-driven optimal control method for an air supply system in proton exchange membrane fuel cells (PEMFCs) is proposed with the aim of improving the PEMFC net output power and operational efficiency. Moreover, a marginal utility-based double-delay deep deterministic policy gradient (MU-4DPG) algorithm is proposed as an offline tuner for the PID controller. The coefficients of the PID controller are rectified and optimized during training in order to enhance the controller's performance. The design of the algorithm draws on the concept of marginal effects in Economics, in that the algorithm continuously switches between different forms of exploration noise during training so as to increase the diversity of samples, improve exploration efficiency and avoid Q-value overfitting, and ultimately improve the robustness of the algorithm. As detailed below, the effectiveness of the control method has been experimentally demonstrated.

Keywords: proton exchange membrane fuel cells, PID controller, data-driven optimal control, double-delay deep deterministic policy gradient, control system

OPEN ACCESS

Edited by:

Yaxing Ren,
University of Warwick,
United Kingdom

Reviewed by:

Chenhao Ren,
The University of Hong Kong, Hong
Kong, SAR China
Du Gang,
North China Electric Power University,
China

*Correspondence:

Jiawen Li
eplijawen@mail.scut.edu.cn

Specialty section:

This article was submitted to
Smart Grids,
a section of the journal
Frontiers in Energy Research

Received: 28 July 2021

Accepted: 02 August 2021

Published: 30 September 2021

Citation:

Li J, Zhu K and Yu T (2021) Data-Driven
Control for Proton Exchange
Membrane Fuel Cells: Method
and Application.
Front. Energy Res. 9:748782.
doi: 10.3389/fenrg.2021.748782

INTRODUCTION

Proton exchange membrane fuel cells (PEMFCs) convert hydrogen energy into electrical energy, and release heat energy directly via an electrochemical reaction (Sun et al., 2019; Yang et al., 2019). The only by-product of the reaction is water, and the whole reaction process does not involve a heat engine process (Yang et al., 2019; Yang et al., 2019) and so it is not limited by the Carnot cycle; thus, the PEMFC is therefore a more environmentally friendly power generation device (Swain and Jena, 2015; Yang et al., 2016; Yang et al., 2017; Yang et al., 2018).

However, PEMFC systems are not entirely compatible with fluctuating load conditions as they have a time lag factored into the air supply system (Yang et al., 2018), and so the supply of air flow does not respond quickly to changes in load. If the air flow in the system is less than that required for the electrochemical reaction, the system cannot be powered properly, leading to severe damage to the proton membrane and reduced working lifetime of the stack (Yang et al., 2020). If the air flow in the system is much higher than the required flow for the electrochemical reaction, the output power is increased, but the power consumption of the compressor increases significantly, which seriously impairs the electrical efficiency of the system. For high-power PEMFC systems, it is necessary to control the air flow so that the air supply system can respond quickly when the load changes (Li et al., 2021; Li and Yu, 2021).

For the regulation of air systems, a large number of control algorithms have been applied in practice.

To achieve optimum performance, it is necessary to co-ordinate the work of the various systems and design the appropriate algorithms for PEMFC control. The most widely used controller in the

industry is the PID controller, which is simple and easy to design, and which is widely used in the PEMFC field. However, PID controllers have certain shortcomings in their ability to resist interference, and researchers of PEMFC systems have explored the use of complex PID control algorithms as potential replacements for the traditional PID controller.

Rgab et al. (2010) proposed a neural network PID controller for controlling the air flow in a PEMFC. Aliasghary (2018) used an interval type II Fuzzy PID controller to control the air flow in a PEMFC system in order to improve the controller's ability to handle uncertainties within the PEMFC system. Baroud et al. (2017) designed a fuzzy adaptive controller that outperforms the fuzzy logic controller and the transcendental PID controller in terms of key performance indicators such as integration squared error, integration time, overshoot of the closed-loop system, stabilization time, and rise time. The fuzzy controller was used by Beiram et al. (2015) to prevent oxygen starvation and ensure maximum net PEMFC power, reducing steady-state errors.

The above innovations can improve the adaptive ability and anti-interference of PID algorithms to a certain extent, but they also increase the calculation volume and calculation time of the algorithms, which makes it harder to ensure the timeliness of the control. In view of these problems, some algorithms that are optimized for the PID algorithm coefficients are more suitable.

The Deep Deterministic Policy Gradient algorithm (DDPG) in deep reinforcement learning is a model-free control algorithm (Zhang et al., 2019; Zhang and Yu, 2019; Zhang et al., 2021) with a simple structure, small computational effort and high robustness, properties which make DDPG an ideal candidate for optimizing the coefficients of PID control algorithms (Zhang et al., 2018). Since DDPG does not require model identification so as to accommodate the uncertainty inherent in nonlinear control, it has been applied in various control fields (Zhang et al., 2016); nevertheless, the algorithm's weak exploration capability explains its low robustness when the algorithm is applied directly as a controller for the control algorithm.

In order to combine the advantages of the DDPG algorithm and PID algorithm, a data-driven method for the optimal control of air flow in a PEMFC is proposed. A marginal utility-based double-delay deep deterministic policy gradient (MU-4DPG) algorithm is proposed as a tuner for the PID controller, one which is trained offline.

The coefficients of the PID controller are rectified and optimized to obtain a PID controller with better performance. The algorithm operates on the principle of marginal effects, a popular analytical concept in the field of Economics, by continuously switching the form of exploration noise in training in order to increase the diversity of samples, improve exploration efficiency and prevent Q-value overfitting, and ultimately improve the robustness of the algorithm to a better performing PID controller.

The innovations in this paper are as follows:

1) A data-driven method for the optimal control of air flow in a PEMFC with proton exchange membrane fuel cells is presented.

2) For this controller, a MU-4DPG algorithm is proposed as the tuner for the PID controller. The algorithm can rectify and optimize the coefficients of the PID controller during offline training to deliver a PID controller with better performance. As mentioned, the design of the algorithm reflects the principle of marginal effects, in that it can enhance the diversity of samples by continuously switching the form of exploration noise in training to improve the exploration efficiency while preventing Q-value overfitting, ultimately improving the robustness of the algorithm and ensuring a better-performing PID controller.

The remainder of this paper comprises the following sections: the PEMFC model is demonstrated in Section *PEMFC Air Supply System Model*, and the algorithm is described in Section *Proposed Method*; the experimental results are analysed and discussed in Section *Case Studies*, and the findings in this paper are summarised in Section *Conclusion*.

PEMFC AIR SUPPLY SYSTEM MODEL

Theoretical Conditions

- 1) The gas settings accord with the Ideal Gas Law.
- 2) The air temperature inside the electrodes is equal to the temperature of the stack.
- 3) When the relative humidity of the gas exceeds 100%, the water vapour concentrates into liquid form.

Air Compressor Model

The air compressor compresses the air and transfers it at a certain pressure to the supply line, where it is cooled and humidified, and then transferred to the fuel cell cathode. The air compressor is the core component within the PEMFC system. The dynamic characteristics of the air compressor drive motor can be expressed by the following set of equations:

$$J_{cp} \frac{d\omega_{cp}}{dt} = \tau_{cm} - \tau_{cp} \quad (1)$$

The motor driving torque of the compressor can be obtained from the static motor equation :

$$\tau_{cm} = \eta_{cm} \frac{k_t}{R_{cm}} (v_{cm} - k_v \omega_{cp}) \quad (2)$$

The load torque of the air compressor can be calculated using the following thermal equation:

$$\tau_{cp} = \frac{C_p T_{atm}}{\omega_{cp} \eta_{cp}} \left[\left(\frac{p_{sm}}{p_{atm}} \right)^{\frac{\gamma-1}{\gamma}} - 1 \right] F_{cp} \quad (3)$$

The temperature of the air leaving the compressor is

$$T_{cp,out} = T_{cp,in} + \frac{T_{cp,in}}{\eta_{cp}} \left[\left(\frac{p_{cp,out}}{p_{cp,in}} \right)^{\gamma-1/\gamma} - 1 \right] \quad (4)$$

The supply of air is completed by the compressor, which consumes electric energy when it is working. Taking into account the power consumed by the compressor and ignoring the power consumption of other equipment in the system, the effective power output of the entire PEMFC system can be expressed as:

$$P_{out} = P_{stack} - P_{com} \quad (5)$$

Return Pipe Model

The principle of the return line is similar to that of the inlet line, but since the temperature of the gas flowing out of the reactor is much lower than that of the gas flowing into the reactor, the effect of temperature is ignored and the principle of the return line can be expressed as follows:

$$\frac{dp_{rm}}{dt} = \frac{R_a T_{rm}}{V_{rm}} (F_{ca,out} - F_{rm,out}) \quad (6)$$

In this paper, according to different pressure ratios, a nonlinear mouth equation is used to determine that $F_{rm,out}$ is

$$F_{rm,out} = \frac{C_{D,rm} A_{T,rm} p_{rm}}{\sqrt{RT_{rm}}} \left(\frac{p_{atm}}{p_{rm}} \right)^{1/\gamma} \sqrt{\frac{2\gamma}{\gamma-1} \left[1 - \left(\frac{p_{atm}}{p_{rm}} \right)^{\gamma-1/\gamma} \right]} \quad (7)$$

$$\frac{p_{atm}}{p_{rm1}} > \left(\frac{2}{\gamma+1} \right)^{\gamma/\gamma-1}$$

Cathode Model

This is obtained from the Ideal Gas Law:

$$\frac{V_{ca}}{RT} \frac{dP_{O_2}}{dt} = w_{O_2,in} - w_{O_2,react} - w_{O_2,out} \quad (8)$$

$$\frac{V_{ca}}{RT} \frac{dP_{N_2}}{dt} = w_{N_2,in} - w_{N_2,out} \quad (9)$$

$$w_{O_2,react} = \frac{NI}{4F} \quad (10)$$

$$w_{O_2,out} = \frac{P_{O_2}}{P_{ca}} w_{ca,out} \quad (11)$$

$$w_{ca,out} = k_{ca} (P_{ca} - P_b) \quad (12)$$

The oxygen excess ratio (OER) is a key variable that significantly influences the fuel cell system's performance. The OER is expressed as follows:

$$\lambda_{O_2} = \frac{W_{O_2,in}}{W_{O_2,ret}} \quad (13)$$

Anode Model

The basic principle of proportional valve control can be expressed as follows:

$$\frac{V_{an}}{RT} \frac{dP_{H_2}}{dt} = w_{H_2,in} - w_{H_2,react} - w_{H_2,out} \quad (14)$$

$$w_{H_2,react} = \frac{NI}{2F} \quad (15)$$

$$w_{H_2,out} = \frac{P_{H_2}}{P_{an}} w_{an,out} \quad (16)$$

$$w_{an,out} = k_{an} (P_{an} - P_b) \quad (17)$$

Supply Pipeline Model

In a PEMFC system, the term "gas supply piping" refers usually to the collection of piping for the cathode only, due to the small size of the anode gas supply piping and return piping. The cathode gas supply pipeline connects the air compressor to the fuel cell reactor cathode and, according to the Law of Conservation of Mass, the settings for gas flowing into the pipeline and the gas flowing out of the pipeline satisfy the following equation:

$$\frac{dp_{sm}}{dt} = \frac{\gamma R_a}{V_{sm}} (F_{cp} T_{cp,out} - F_{sm,out} T_{sm}) \quad (18)$$

When the high-temperature gas from the compressor enters the gas supply line, its temperature and pressure change, so, according to the Law of Energy Conservation and the Ideal Gas Equation, the change of pressure in the gas supply line can be expressed as follows:

$$F_{sm,out} = k_{sm,out} (p_{sm} - p_{ca}) \quad (19)$$

PROPOSED METHOD

Technical Principles

The control model includes the PEMFC reactor, the air compressor and its PID controller. The controller of the PID is equated to an agent undergoing training, which in turn enables the agent to adapt to the non-linear characteristics of the PEMFC and improve the overall control performance of the OER and output voltage. For pre-learning, the proposed algorithm is used to rectify the coefficients k_p , k_i , and k_d of the PID controller.

The tuning methods employ a linear quadratic Gaussian (LQG) objective function.

In order to maintain the water level while minimizing control effort u , the controllers employ the following LQG criterion:

$$J = \lim_{T \rightarrow \infty} E \left(\frac{1}{T} \int_0^T ((y_{ref} - y)^2(t) + 0.01u^2(t)) dt \right) \quad (20)$$

When applied online, the agent issues the optimal voltage of the air compressor once it has received data on the status of the PEMFC. The control interval of the agent is 0.01 s.

DDPG

The DDPG algorithm is a policy-based RL method which is based on an empirical replay approach and which uses deep neural networks as nonlinear function approximators to construct $Q(s, a; \theta^Q)$, $\mu(s; \theta^\mu)$ and $\mu'(s; \theta^\mu)$ critic network model action and target Q values.

$$\mu'(s_t) = \mu(s|\theta^\mu) + N \quad (21)$$

The critic network is updated by minimizing the loss function. The $L(\theta^Q)$ network parameters are shown as follows:

$$y_t = r(s_t, a_t) + \gamma Q'(s_{t+1}, \mu'(s_{t+1}|\theta^\mu)\theta^Q) \quad (22)$$

The gradient of the actor network, which depends on which the actor network parameters are updated, is expressed as follows:

$$\nabla_{\theta^\mu} J(\theta^\mu) = \frac{1}{N} \sum_i \nabla_a Q(s, a|\theta^Q) \Big|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu} \mu(s|\theta^\mu) \Big|_{s_i} \quad (23)$$

DDPG improves the stability of the learning process by slowly updating the weights of the actor target network and the critic target network:

$$\begin{cases} \theta^\mu \leftarrow \tau\theta + (1 - \tau)\theta' \\ \theta^Q \leftarrow \tau\theta + (1 - \tau)\theta' \end{cases} \quad (24)$$

MU-4DPG Algorithm

MU-4DPG is an extension of the DDPG algorithm, which in turn has low robustness due to its single exploration principle. This problem, similar to the concept of marginal effect in Economics, is due to the fact that the exploration principle is too homogeneous, and so the agent receives only a fixed number of types of actions, and struggles to obtain richer samples. Marginal effects are defined as those in which a successive increase in one input, when other inputs are fixed, results in a gradual decrease in the benefit, that is, the amount of output per unit of input added decreases when the added input exceeds a certain level. Therefore, the diversity of samples obtained by using the same noise model over and over again decreases with an increase of episodes. In order to solve this problem, the method proposed in this paper improves the diversity of samples by continuously switching between different forms of exploration noise during training, improving the exploration efficiency, and adopting a number of strategies in order to prevent Q-value overfitting. The result is an algorithm with higher robustness, and a better-performing PID controller.

The MU-4DPG algorithm includes two critics and one actor, and, in order to solve the Q overestimation problem that occurs in the DDPG algorithm, the proposed algorithm uses three techniques: clipping multiple Q learning, delayed policy update, and smooth regularization of the target policy.

- 1) Clipped multi-Q learning. The “student” agent in MU-4DPG uses clipped multi-Q learning to calculate the target value:

$$y_t^1 = r(s_t, a_t) + \gamma \min_{i=1,2,3} Q_{\theta^i}(s_{t+1}, \pi_{\phi_1}(s_{t+1})) \quad (25)$$

- 2) Policy delayed update. After every d update of the critic network, an update of the actor network is performed so as to ensure that the actor network can be updated with a low Q-value error, in order to improve the update efficiency of the actor network.

- 3) Target policy smoothing regularization. The algorithm introduces a regularization method to reduce the variance

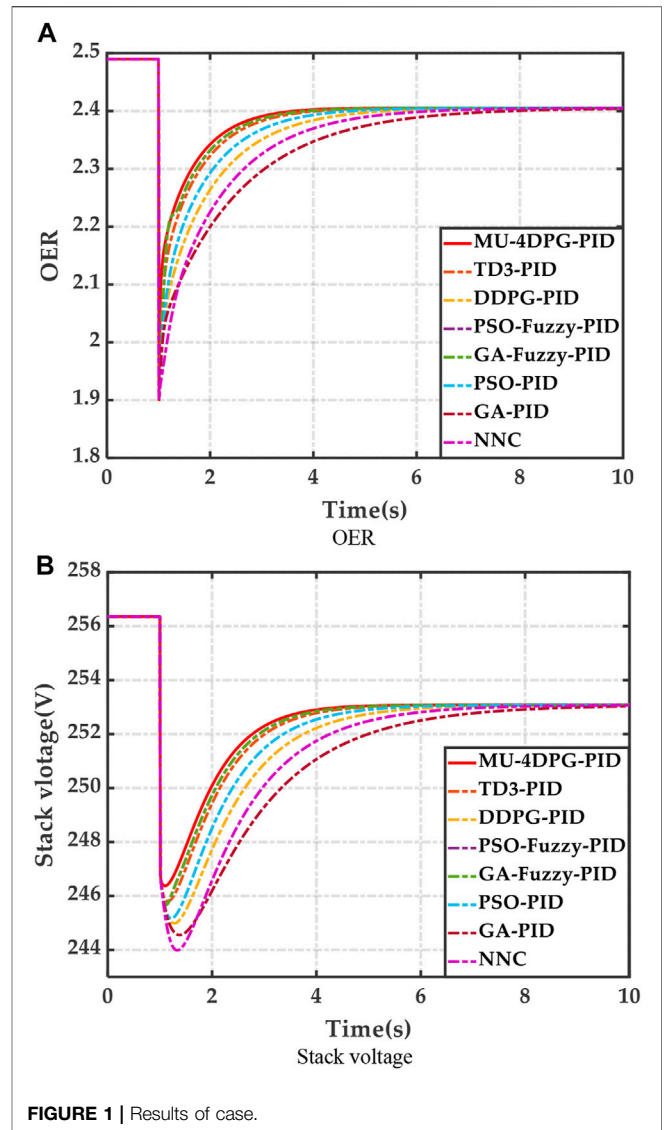


FIGURE 1 | Results of case.

of the target values by smoothing the Q-estimates by bootstrapping estimates of similar state action pairs:

$$y_t = r(s_t, a_t) + E_\epsilon [Q_{\theta'}(s_{t+1}, \pi_{\phi'}(s_{t+1}) + \epsilon)] \quad (26)$$

Smooth regularization is also achieved by adding a random noise to the target strategy and averaging over the mini-batch:

$$y_t = r(s_t, a_t) + \gamma \min_{i=1,2} Q_{\theta_i}(s_{t+1}, \pi_{\phi'}(s_{t+1}) + \epsilon) \quad (27)$$

$$\epsilon \sim \text{clip}(N(0, \sigma), -c, c) \quad (28)$$

- 4) Marginal effect mechanism. The actor network in the agent switches between Gaussian noise and OU noise, depending on the episodes, in order to obtain a more diverse sample.

Episodes are 0–1,000, 2000–3,000, 4,000–5,000, 7,000–8,000, 9,000–10,000 when OU noise is used for exploration:

$$a_{OU}^j = \pi_{\phi}^j(s) + N_{OU}^j \quad (29)$$

whereby $\pi_{\phi}^j(s)$ is the actor network policy, and N_{OU} is the OU noise.

Episodes are 1,001–1999, 3,001–3,999, 5,001–5,999, 6,001–6,999, 8,001–8,999 when explored using Gaussian noise:

$$a_{Gaussian}^m = \pi_{\theta}^m(s) + \mathcal{N}_{Gaussian}^m \quad (30)$$

whereby $\pi_{\theta}^m(s)$ is the actor network policy, and $N_{Gaussian}$ is Gaussian noise.

CASE STUDIES

In the simulation, the MU-4DPG tuning PID (MU-4DPG-PID) was compared against the TD3 tuning PID (TD3-PID) controller, DDPG tuning PID (DDPG-PID) controller, PSO optimized fuzzy PID controller (PSO-Fuzzy-PID), GA-optimized fuzzy PID controller (GA-Fuzzy-PID), PSO-optimized PID controller (PSO-PID), GA-optimized PID controller (GA-PID), and neural network control (NNC). The results are shown in **Figures 1A,B**.

- 1) Comparison of the proposed algorithm with other algorithms. According to **Figures 1A,B**, the OER of the MU-4DPG algorithm has a better climbing speed response, a more stable OER, and a smaller overshoot. In addition, the MU-4DPG algorithm has a more rapid response time for the stack voltage and a smaller overshoot, and it does not permit large fluctuations. By contrast, the other algorithms each have a large overshoot in both the OER and output voltage, and a slower response time, resulting in large oscillations in both the OER and the stack voltage, which can lead to large oscillations in the output voltage. Therefore, the MU-4DPG algorithm has the best control performance.
- 2) The reasons for this phenomenon are: the robustness of the algorithm is reduced compared to other DRL algorithms because more techniques are not used in pre-learning, resulting in a large output voltage overshoot and output voltage fluctuations that affect the output performance of the PEMFC. the Fuzzy-based algorithm is too simple in its rules, resulting in low robustness and adaptive capability of this class of algorithms. The PSO-PID, and GA-PID algorithms inside the conventional control algorithm do not have the adaptive capability to adjust the PID parameters, and therefore have difficulty in adapting to the non-linearity of the PEMFC. The NNC algorithm, on the other hand, relies on the effects of training, resulting in low

robustness and therefore the lowest control performance of this class of algorithms.

CONCLUSION

- 1) This paper presents a large-scale deep reinforcement learning based adaptive optimal PID controller for controlling proton exchange membrane fuel cell (PEMFC) air flow.
- 2) A marginal utility-based double-delay deep deterministic policy gradient (MU-4DPG) algorithm is proposed as a tuner for the PID controller. The coefficients of the PID controller are rectified and optimized during off-line training to obtain a PID controller with better performance and fixed coefficients. The algorithm operates on the Economics principle of marginal effects, in that it continuously switches between different forms of exploration noise in training in order to increase the diversity of samples, improve exploration efficiency and prevent Q-value overfitting, which altogether lead to an improvement in the robustness of the algorithm and a better-performing PID controller.
- 3) The results of the simulation involving both the proposed controller and two groups of existing controllers (adaptive PID and conventional control) indicate that the MU-4DPG-PID controller is able to maintain a stable output voltage while effectively avoiding oxygen starvation or oxygen supersaturation in the fuel cell.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JL: Conceptualization, Methodology, Software. Zhang: Writing-Reviewing and Editing. Data curation, Writing-Original draft preparation. Visualization, Investigation. KZ: Validation. TY: Supervision.

FUNDING

This work was jointly supported by National Natural Science Foundation of China (U2066212).

REFERENCES

Aliasghary, M. (2018). Control of PEM Fuel Cell Systems Using Interval Type-2 Fuzzy PID Approach. *Fuel Cells* 18, 449–456. doi:10.1002/fuce.201700157

Baroud, Z., Benmiloud, M., Benalia, A., and Ocampo-Martinez, C. (2017). Novel Hybrid Fuzzy-PID Control Scheme for Air Supply in PEM Fuel-Cell-Based Systems. *Int. J. Hydrogen Energ.* 42, 10435–10447. doi:10.1016/j.ijhydene.2017.01.014

Beirami, H., Shabestari, A. Z., and Zerafat, M. M. (2015). Optimal PID Plus Fuzzy Controller Design for a PEM Fuel Cell Air Feed System Using the Self-Adaptive

- Differential Evolution Algorithm. *Int. J. Hydrogen Energ.* 40, 9422–9434. doi:10.1016/j.ijhydene.2015.05.114
- Li, J., and Yu, T. (2021). A New Adaptive Controller Based on Distributed Deep Reinforcement Learning for PEMFC Air Supply System. *Energ. Rep.* 7, 1267–1279. doi:10.1016/j.egy.2021.02.043
- Li, J., Yu, T., Zhang, X., Li, F., Lin, D., and Zhu, H. (2021). Efficient Experience Replay Based Deep Deterministic Policy Gradient for AGC Dispatch in Integrated Energy System. *Appl. Energ.* 285, 116386. doi:10.1016/j.apenergy.2020.116386
- Rgab, O., Yu, D. L., and Gomm, J. B. (2010). Polymer Electrolyte Membrane Fuel Cell Control with Feed-Forward and Feedback Strategy. *Int. J. Eng. Sci. Technol.* 2, 80–91. doi:10.4314/ijest.v2i10.64012
- Sun, L., Jin, Y., Pan, L., Shen, J., and Lee, K. Y. (2019). Efficiency Analysis and Control of a Grid-Connected PEM Fuel Cell in Distributed Generation. *Energ. Convers. Manage.* 195, 587–596. doi:10.1016/j.enconman.2019.04.041
- Swain, P., and Jena, D. (2015). “PID Control Design for the Pressure Regulation of PEM Fuel Cell,” in 2015 International Conference on Recent Developments in Control, Automation and Power Engineering (RDCAPE), Noida, India, 12–13 March 2015 (IEEE), 286–291. doi:10.1109/rdcape.2015.7281411
- Yang, B., Jiang, L., Wang, L., Yao, W., and Wu, Q. H. (2016). Nonlinear Maximum Power Point Tracking Control and Modal Analysis of DFIG Based Wind Turbine. *Int. J. Electr. Power Energ. Syst.* 74, 429–436. doi:10.1016/j.jepes.2015.07.036
- Yang, B., Zhang, X., Yu, T., Shu, H., and Fang, Z. (2017). Grouped Grey Wolf Optimizer for Maximum Power point Tracking of Doubly-Fed Induction Generator Based Wind Turbine. *Energ. Convers. Manage.* 133, 427–443. doi:10.1016/j.enconman.2016.10.062
- Yang, B., Yu, T., Shu, H., Dong, J., and Jiang, L. (2018). Robust Sliding-Mode Control of Wind Energy Conversion Systems for Optimal Power Extraction via Nonlinear Perturbation Observers. *Appl. Energ.* 210, 711–723. doi:10.1016/j.apenergy.2017.08.027
- Yang, B., Yu, T., Shu, H., Han, Y., Cao, P., and Jiang, L. (2019a). Adaptive Fractional-Order PID Control of PMSG-Based Wind Energy Conversion System for MPPT Using Linear Observers. *Int. Trans. Electr. Energy Syst.* 29 (9), e2697. doi:10.1002/etep.2697
- Yang, B., Zhong, L., Zhang, X., Shu, H., Yu, T., Li, H., et al. (2019b). Novel Bio-Inspired Memetic Salp Swarm Algorithm and Application to MPPT for PV Systems Considering Partial Shading Condition. *J. Clean. Prod.* 215, 1203–1222. doi:10.1016/j.jclepro.2019.01.150
- Yang, B., Yu, T., Zhang, X., Li, H., Shu, H., Sang, Y., et al. (2019c). Dynamic Leader Based Collective Intelligence for Maximum Power point Tracking of PV Systems Affected by Partial Shading Condition. *Energ. Convers. Manage.* 179 (18), 286–303. doi:10.1016/j.enconman.2018.10.074
- Yang, B., Wang, J., Zhang, X., Yu, T., Yao, W., Shu, H., et al. (2020). Comprehensive Overview of Meta-Heuristic Algorithm Applications on Pv Cell Parameter Identification. *Energ. Convers. Manage.* 208, 112595. doi:10.1016/j.enconman.2020.112595
- Zhang, X., and Yu, T. (2019). Fast Stackelberg Equilibrium Learning for Real-Time Coordinated Energy Control of a Multi-Area Integrated Energy System. *Appl. Therm. Eng.* 153, 225–241. doi:10.1016/j.applthermaleng.2019.02.053
- Zhang, X., Xu, H., Yu, T., Yang, B., and Xu, M. (2016). Robust Collaborative Consensus Algorithm for Decentralized Economic Dispatch with a Practical Communication Network. *Electric Power Syst. Res.* 140, 597–610. doi:10.1016/j.epsr.2016.05.014
- Zhang, X., Yu, T., Xu, Z., and Fan, Z. (2018). A Cyber-Physical-Social System with Parallel Learning for Distributed Energy Management of a Microgrid. *Energy* 165, 205–221. doi:10.1016/j.energy.2018.09.069
- Zhang, X., Li, S., He, T., Yang, B., Yu, T., Li, H., et al. (2019). Memetic Reinforcement Learning Based Maximum Power point Tracking Design for PV Systems under Partial Shading Condition. *Energy* 174, 1079–1090. doi:10.1016/j.energy.2019.03.053
- Zhang, X., Tan, T., Zhou, B., Yu, T., Yang, B., and Huang, X. (2021). Adaptive Distributed Auction-Based Algorithm for Optimal Mileage Based AGC Dispatch with High Participation of Renewable Energy. *Int. J. Electr. Power Energ. Syst.* 124, 106371. doi:10.1016/j.jepes.2020.106371

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Li, Zhu and Yu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.