A New Self-Learning Optimization Management Scheme for Smart Grid Systems

Qinglai Wei and Zehua Liao

Abstract—**This paper presents a new energy optimization management scheme for the smart grid system (SGS), which considers not only the inclusion of renewable energy in energy management but also the feedback sale of redundant renewable energy to the grid, so as to obtain better optimization management results and avoid energy waste. A new performance index function is constructed, which aims to save users' electricity costs and protect energy storage batteries. We construct a periodic iterative self-learning method based on the characteristics that the relevant data of SGS are approximately periodic. The proposed self-learning optimization method is based on adaptive dynamic programming (ADP). The initial iterative control law is obtained by pre-training, and the designed method can be realized by neural networks. Simulation experiments are carried out by using the relevant data of SGS for four weeks. The experimental results show that the new energy management scheme can help users save a lot of electricity costs, and is more effective than some common methods.**

***Index Terms*—Smart grid system, energy management, renewable energy, self-learning, adaptive dynamic programming**

I. Introduction

Smart grid system (SGS) is an evolution of the present power grid structures due to the developing demand of energy, the growing charge of fossil fuels, and the improvement of new and progressive intelligent sciences [1–3]. However, many researches on smart grid focus on a small part of energy management and limited control actions [4, 5]. In this paper, a new SGS is proposed, which is composed of power resources, load demand, and energy storage equipments. Distributed energy resources are considered in the SGS with a variety of energy management approaches. The power resources contain the power grid and the renewable resources, which include solar energy, wind energy, and so on [6, 7]. This SGS gives full play to the advantages of clean, low pollution, and low price of renewable energy, and at the same time, considers selling the surplus renewable energy to the power grid to solve the problem of possible renewable energy waste.

The focus of SGS optimization is the design of its controller. However, it is difficult to build a specific SGS model, and some traditional control methods are not effective. Adaptive dynamic programming (ADP) [8, 9] is an effective technique to solve optimal control problems for nonlinear systems such as SGS [10–12]. At present, there have been many research results of ADP method in solving optimal control problems of SGS, and the corresponding optimal control results have been obtained [13–16]. However, as mentioned before, these studies usually take into account the relatively limited types of energy and limited control actions. In the work of Ref. [17], the idea of selling excess renewable sources was considered, and the higher economic benefits than before were obtained. However, limited control actions are considered in Ref. [17], which cannot prove the convergence and optimality of its algorithm.

In this paper, a novel smart grid system with a variety of energy management actions is considered, and a new self-learning dual ADP method is developed for obtaining the optimal control. The numerical results and the comparisons with the methods proposed by Refs. [17–19] will illustrate the effectiveness and superiority of the present dual ADP method.

The rest of this paper is as follows. Section II is the problem formulation of the studied content. Section III develops the optimal control method for the SGS. Section IV shows the numerical experiments and comparison experiments. Finally, Section V concludes this paper.

II. Problem Formulation

The structure of the SGS is shown in Fig. 1, which also shows the power flows of the control model. In this situation, how to optimize the control law of the energy management system is the key to saving the cost of the SGS.

For the grid side, there are three possible power flows: (i) Power grid directly supplies power to meet the load demand. (ii) It can charge the battery which does not supply the load. (iii) The power grid supplies the load and simultaneously charges the battery. For the load side, the load demand is generally determined by the users, and is required to satisfy. There are also three possible power flows to meet the load demand: (i) The load can receive the power from the power grid. (ii) It can also receive the power from the energy storage equipment. (iii) It can receive the power both from the grid and the energy storage equipment. For the renewable energy side, there are also three possible power flows: (i) The renewable energy directly supplies power to meet the load. (ii) It can charge the battery which does not supply the load. (iii) The excess energy can sell to the grid. In this situation, the energy management system is a suitable controller to optimize the energy management for the SGS. The main control modes

Manuscript received: 15 September 2023; revised: 24 December 2023; accepted: 8 April 2024. *(Corresponding author: Qinglai Wei.)*

Citation: Q. Wei and Z. Liao, A new self-learning optimization management scheme for smart grid systems, *IJICS*, 2024, 29(2), 66–70.

Qinglai Wei and Zehua Liao are with School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China, with State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and with Institute of Systems Engineering, Macau University of Science and Technology, Macao 999078, China (e-mail: qinglai.wei@ia.ac.cn; liaozehua2017@ia.ac.cn).

Digital Object Identifier 10.62678/IJICS202403.10108

of the energy management system include idle mode, discharging mode, and charging mode.

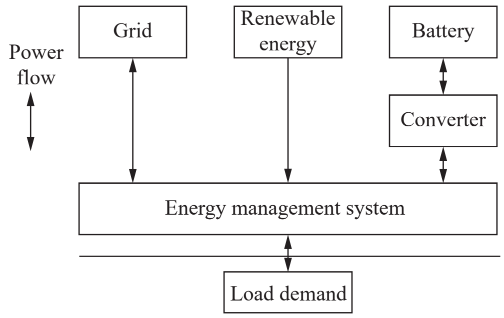

Figure 1 Structure of SGS.

Based on the above analysis, we have $E_{\mathrm{g}(\sigma)} = E_{\mathrm{gl}(\sigma)}+E_{\mathrm{gb}(\sigma)}-E_{\mathrm{rg}(\sigma)}$, $E_{\mathrm{r}(\sigma)} = E_{\mathrm{rl}(\sigma)}+E_{\mathrm{rb}(\sigma)}+E_{\mathrm{rg}(\sigma)}$, and $E_{\mathrm{b}(\sigma)} = E_{\mathrm{bl}(\sigma)}-E_{\mathrm{gb}(\sigma)}-E_{\mathrm{rb}(\sigma)}$, where σ is the time index, l, b, g, and r denote load, battery, grid, and renewable energy, respectively. $E_{\mathrm{bl}(\sigma)}$ is used to denote the energy from the battery to the load, and other notations have similar meanings.

For $\sigma = 0,1,2,...$, the load balance is defined as

$$E_{\mathrm{l}(\sigma)} = E_{\mathrm{bl}(\sigma)}+E_{\mathrm{gl}(\sigma)}+E_{\mathrm{rl}(\sigma)} \tag{1}$$

According to the power flows of the distributed energy sides, the new relations are defined as

$$\mathcal{E}_{\mathrm{r}(\sigma)} = \begin{cases} E_{\mathrm{r}(\sigma)}-E_{1(\sigma)},\ E_{1(\sigma)}-E_{\mathrm{r}(\sigma)} \geqslant 0; \\ 0,\ E_{1(\sigma)}-E_{\mathrm{r}(\sigma)} < 0 \end{cases} \tag{2}$$

and

$$\mathcal{E}_{\mathrm{l}(\sigma)} = \begin{cases} E_{\mathrm{l}(\sigma)}-E_{\mathrm{r}(\sigma)},\ E_{\mathrm{l}(\sigma)}-E_{\mathrm{r}(\sigma)} \geqslant 0; \\ 0,\ E_{\mathrm{l}(\sigma)}-E_{\mathrm{r}(\sigma)} < 0 \end{cases} \tag{3}$$

Define the battery energy $B_{\mathrm{b}(\sigma)}$ as

$$\begin{aligned} B_{\mathrm{b}(\sigma+1)} = {} & B_{\mathrm{b}(\sigma)}-E_{\mathrm{bl}(\sigma)}\times\eta(E_{\mathrm{bl}(\sigma)})+ \\ & (E_{\mathrm{rb}(\sigma)}+E_{\mathrm{gb}(\sigma)})\times\eta(E_{\mathrm{rb}(\sigma)}+E_{\mathrm{gb}(\sigma)}) \end{aligned} \tag{4}$$

where $B_{\mathrm{b}(\sigma)}$ should satisfy $B_{\mathrm{b,min}} \leqslant B_{\mathrm{b}(\sigma)} \leqslant B_{\mathrm{b,max}}$ to guarantee the security of SGS. The battery efficiency is defined as $\eta(E_{(\cdot)}) = 0.898-0.173|E_{(\cdot)}|/E_{\mathrm{rate}}$ according to Refs. [20, 21], where E_{rate} represents the rated energy of the battery or the rated charge/discharge rate of the battery.

Define $s_{1(\sigma)} = E_{\mathrm{gl}(\sigma)}+E_{\mathrm{gb}(\sigma)}-mE_{\mathrm{rg}(\sigma)}$, $s_{2(\sigma)} = B_{\mathrm{b}(\sigma)}-B^{o}_{\mathrm{b}}$, $v_{1(\sigma)} = E_{\mathrm{bl}(\sigma)}-E_{\mathrm{gb}(\sigma)}$, and $v_{2(\sigma)} = E_{\mathrm{rb}(\sigma)}$, where m represents the ratio of electricity from renewable energy to grid to the total electricity from renewable energy, and subscript o indicates the nominal capacity of the battery. Let $s_{(\sigma)} = [s_{1(\sigma)}, s_{2(\sigma)}]^{\mathrm{T}}$ and $v_{(\sigma)} = [v_{1(\sigma)}, v_{2(\sigma)}]^{\mathrm{T}}$, the discrete system of SGS can be defined as

$$\begin{aligned} s_{(\sigma+1)} = {} & \Lambda(s_{(\sigma)}, v_{(\sigma)}, \sigma) = \\ & \begin{pmatrix} \mathcal{E}_{\mathrm{l}(\sigma)}-v_{1(\sigma)}-m(\mathcal{E}_{\mathrm{r}(\sigma)}-v_{2(\sigma)}) \\ s_{2(\sigma)}-(v_{1(\sigma)}-v_{2(\sigma)})\eta(v_{1(\sigma)}-v_{2(\sigma)}) \end{pmatrix} \end{aligned} \tag{5}$$

where the utility function $\Lambda(s_{(\sigma)}, v_{(\sigma)}, \sigma)$ is defined as

$$\begin{aligned} & \Lambda(s_{(\sigma)}, v_{(\sigma)}, \sigma) = \\ & \Big(r_1\big(C_{(\sigma)}(E_{\mathrm{gl}(\sigma)}+E_{\mathrm{gb}(\sigma)})-\bar{C}_{(\sigma)}E_{\mathrm{rg}(\sigma)}\big)^2+ \\ & r_2\big(B_{\mathrm{b}(\sigma)}-B^{o}_{\mathrm{b}}\big)^2+r_3\big(E_{\mathrm{bl}(\sigma)}-E_{\mathrm{gb}(\sigma)}\big)^2+r_4\big(E_{\mathrm{rb}(\sigma)}\big)^2\Big) = \\ & \Big(r_1\big(C_{(\sigma)}\big(E_{\mathrm{gl}(\sigma)}+E_{\mathrm{gb}(\sigma)}-mE_{\mathrm{rg}(\sigma)}\big)\big)^2+ \\ & r_2\big(B_{\mathrm{b}(\sigma)}-B^{o}_{\mathrm{b}}\big)^2+r_3\big(E_{\mathrm{bl}(\sigma)}-E_{\mathrm{gb}(\sigma)}\big)^2+r_4\big(E_{\mathrm{rb}(\sigma)}\big)^2\Big) \end{aligned} \tag{6}$$

where r_1, r_2, r_3, r_4, and m are given positive constants, $C_{(\sigma)}$ and $\bar{C}_{(\sigma)} = mC_{(\sigma)}$ are electricity price and the price of selling power, respectively.

The optimal performance index function $J^*(s_{(\sigma)}, \sigma)$ satisfies the Bellman equation

$$J^*(s_{(\sigma)}, \sigma) = \inf_{v_{(\sigma)}}\{\Lambda(s_{(\sigma)}, v_{(\sigma)}, \sigma)+\gamma J(s_{(\sigma+1)}, \sigma+1)\} \tag{7}$$

where $*$ represents the function under optimal control, and γ represents the discount factor used to weigh the current and future costs in the Bellman equation.

III. Method

Define a new value function as

$$\min_{v_{(\sigma)}} \mathcal{V}^*[s_{(\sigma)}, v_{(\sigma)}, \sigma] = J^*[s_{(\sigma)}, \sigma] \tag{8}$$

Assume that energy sources are all periodic functions with the period of Θ. Then, define the accumulated cost during the periodic iteration process as

$$\Upsilon[s_{(\varsigma)}, \Xi_{(\varsigma)}] = \sum_{\varpi=0}^{\Theta-1}\gamma^{\varpi}\Lambda[s_{(\varsigma+\varpi)}, v_{(\varsigma+\varpi)}, \varpi] \tag{9}$$

where ς represents the iteration variable, $s_{(\varsigma)}$ represents the system state vector, $\Xi_{(\varsigma)} = \{v_{(\varsigma)}, v_{(\varsigma+1)}, \ldots, v_{(\varsigma+\Theta-1)}\}$ represents the control sequence, $v_{(\varsigma)}$ represents the control input vector, and ϖ denotes the index of the control sequence, respectively. Hence, Eq. (8) can be expressed as

$$\begin{aligned} \mathcal{V}^*[s_{(\varsigma)}, \Xi_{(\varsigma)}] = {} & \Upsilon[s_{(\varsigma)}, \Xi_{(\varsigma)}]+ \\ & \gamma^{\Theta}\min_{\Xi_{(\varsigma+\Theta)}}\mathcal{V}^*[s_{(\varsigma+\Theta)}, \Xi_{(\varsigma+\Theta)}] \end{aligned} \tag{10}$$

The optimal control is derived by

$$\Xi^*_{(\varsigma)} = \arg\min_{\Xi_{(\varsigma)}}\mathcal{V}^*[s_{(\varsigma)}, \Xi_{(\varsigma)}] \tag{11}$$

Next, a novel ADP method is developed with dual iteration processes. In the external iteration, policy iteration is used to gain the optimal iterative value function. Define $\Xi_0(s_{(\varsigma)})$ as the initial control that the particle swarm optimization (PSO) method is adopted to achieve $\Xi_0(s_{(\varsigma)})$. Implement the PSO method, the sequence $\{p_{\mathrm{g}(0)}, p_{\mathrm{g}(1)}, \ldots, p_{\mathrm{g}(D-1)}\}$ is obtained, where D is the length of the sequence, and p_{g} is the position of the particle in PSO and also refers to the control $v(\sigma)$ in Eq. (5). Thus, we have $\Xi_0(s_{(\varsigma)}) = \{p_{\mathrm{g}(0)}, p_{\mathrm{g}(1)}, \ldots, p_{\mathrm{g}(D-1)}\}$. Then, the initial value function is obtained as

$$\begin{aligned} \mathcal{V}_0[s_{(\varsigma)}, \Xi_{(\varsigma)}] = {} & \Upsilon[s_{(\varsigma)}, \Xi_{(\varsigma)}]+ \\ & \gamma^{\Theta}\mathcal{V}_0[s_{(\varsigma+\Theta)}, \Xi_0(s_{(\varsigma+\Theta)})] \end{aligned} \tag{12}$$

Next, for number of iterative step $\epsilon = 1,2,\ldots$, the iteration is proceeded by

$$\Xi_\epsilon(s_{(\varsigma)}) = \arg\min_{\Xi_{(\varsigma)}} \mathcal{V}_{\epsilon-1}[s_{(\varsigma)}, \Xi_{(\varsigma)}] \tag{13}$$

and

$$\mathcal{V}_\epsilon[s_{(\varsigma)}, \Xi_{(\varsigma)}] = \Upsilon[s_{(\varsigma)}, \Xi_{(\varsigma)}] + \gamma^{\Theta}\mathcal{V}_\epsilon[s_{(\varsigma+\Theta)}, \Xi_\epsilon(s_{(\varsigma+\Theta)})] \tag{14}$$

In the internal iteration, value iteration is used to achieve the control in Eq. (13). The initial iterative value function is defined as

$$\mathcal{V}_0^0[s_{(\varsigma)}, \nu_{(\varsigma)}] = \mathcal{V}_0[s_{(\varsigma)}, \Xi_{(\varsigma)}] \tag{15}$$

The corresponding control $\nu_0^0(s_{(\varsigma)})$ is

$$\nu_0^0(s_{(\varsigma)}) = \arg\min_{\nu_{(\varsigma)}} \mathcal{V}_0^0[s_{(\varsigma)}, \nu_{(\varsigma)}] \tag{16}$$

For $\epsilon = 0$ and discrete moment $\varepsilon = 1, 2, \ldots, \Theta$, the iteration is proceeded by

$$\begin{aligned} &\mathcal{V}_0^\varepsilon[s_{(\varsigma)}, \nu_{(\varsigma)}] = \\ &\Lambda[s_{(\varsigma)}, \nu_{(\varsigma)}, \varepsilon-1] + \gamma \min_{\nu_{(\varsigma+1)}} \mathcal{V}_0^{\varepsilon-1}[s_{(\varsigma+1)}, \nu_{(\varsigma+1)}] = \\ &\Lambda[s_{(\varsigma)}, \nu_{(\varsigma)}, \varepsilon-1] + \gamma \mathcal{V}_0^{\varepsilon-1}\left[s_{(\varsigma+1)}, \nu_0^{\varepsilon-1}(s_{(\varsigma+1)})\right] \end{aligned} \tag{17}$$

and

$$\nu_0^\varepsilon(s_{(\varsigma)}) = \arg\min_{\nu_{(\varsigma)}} \mathcal{V}_0^\varepsilon[s_{(\varsigma)}, \nu_{(\varsigma)}] \tag{18}$$

where

$$s_{(\varsigma+1)} = \begin{pmatrix} \mathcal{E}_{l(\Theta-\varepsilon)} - \nu_{1(\varsigma)} - m(\mathcal{E}_{r(\Theta-\varepsilon)} - \nu_{2(\varsigma)}) \\ s_{2(\varsigma)} - (\nu_{1(\varsigma)} - \nu_{2(\varsigma)})\eta(\nu_{1(\varsigma)} - \nu_{2(\varsigma)}) \end{pmatrix} \tag{19}$$

For $\epsilon = 1, 2, \ldots$, let $\mathcal{V}_\epsilon^0[s_{(\varsigma)}, \nu_{(\varsigma)}] = \mathcal{V}_{\epsilon-1}^{\Theta}[s_{(\varsigma)}, \Xi_{(\varsigma)}]$, and

$$\nu_\epsilon^0(s_{(\varsigma)}) = \arg\min_{\nu_{(\varsigma)}} \mathcal{V}_\epsilon^0[s_{(\varsigma)}, \nu_{(\varsigma)}] \tag{20}$$

For $\varepsilon = 1, 2, \ldots, \Theta$, the iteration is proceeded by

$$\begin{aligned} &\mathcal{V}_\epsilon^\varepsilon[s_{(\varsigma)}, \nu_{(\varsigma)}] = \\ &\Lambda[s_{(\varsigma)}, \nu_{(\varsigma)}, \varepsilon-1] + \gamma \mathcal{V}_\epsilon^{\varepsilon-1}\left[s_{(\varsigma+1)}, \nu_\epsilon^{\varepsilon-1}(s_{(\varsigma+1)})\right] \end{aligned} \tag{21}$$

and

$$\nu_\epsilon^\varepsilon(s_{(\varsigma)}) = \arg\min_{\nu_{(\varsigma)}} \mathcal{V}_\epsilon^\varepsilon[s_{(\varsigma)}, \nu_{(\varsigma)}] \tag{22}$$

Hence, the control in Eq. (13) can be solved by

$$\Xi_\epsilon(s_{(\varsigma)}) = \left\{\nu_\epsilon^{\Theta-1}(s_{(\varsigma)}), \nu_\epsilon^{\Theta-2}(s_{(\varsigma)}), \ldots, \nu_\epsilon^0(s_{(\varsigma)})\right\} \tag{23}$$

which is used to guide the dispatching of distributed energy resources.

In this paper, neural networks can be utilized to implement the dual ADP method. We use the critic and action networks with the three-layer back propagation (BP) structure to approximate the value function in Eq. (10) and control in Eq. (11), respectively.

IV. Numerical Analysis

We set the simulation period as four weeks (672 h). Let $B_{b,max} = 80$ kWh, $B_{b,min} = 20$ kWh, and $B_{b(0)} = 60$ kWh. Let $m = 0.5$, $r_1 = 1.0$, $r_2 = 0.4$, $r_3 = 0.1$, and $r_4 = 0.1$. Let $\gamma = 0.995$. The data of energy resources and electricity price are shown in Fig. 2.

The developed dual ADP method is implemented for enough iterations and the optimal scheduling results are obtained, as shown in Fig. 3. Figure 3 shows the optimal scheduling actions of the distributed energy resources at each time, including renewable energy distribution and battery charging (or discharging) control. Generally, energy from the power grid is used when the electricity price is low, while renewable energy and energy stored in the battery are utilized

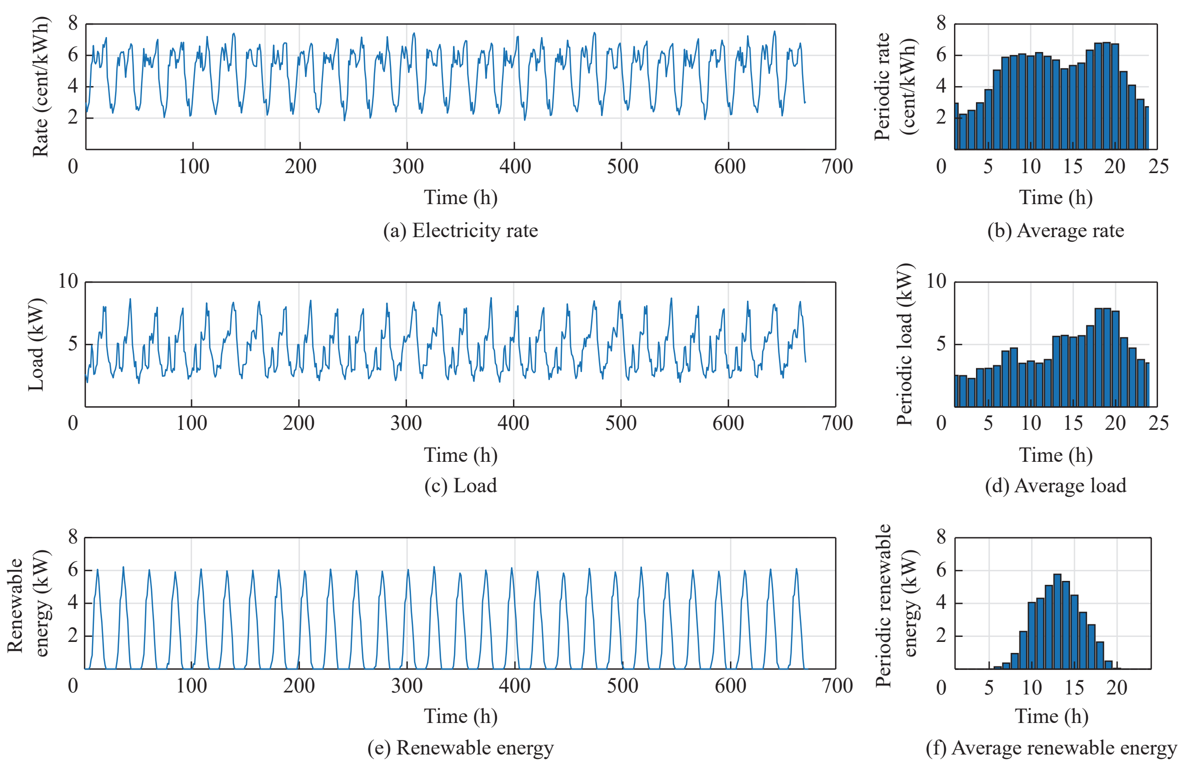

Figure 2 Change of energy price and load demand in smart grid.

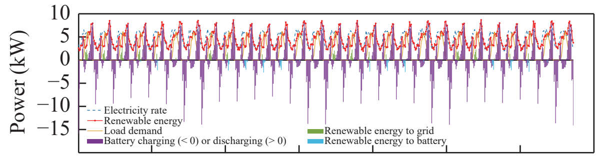

Figure 3 Optimal scheduling of SGS in four weeks.

when the price is high. The change of residual energy of the battery is shown in Fig. 4, which shows the power flow process more clearly. The results verify the effectiveness of the developed method.

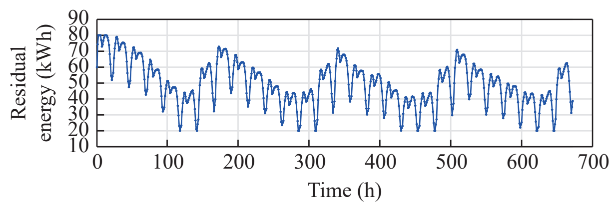

Figure 4 Energy change of the battery.

The adaptive simulated annealing genetic algorithm (ASAGA) method [19], model predictive control (MPC) method [20], and time-based quantum leap (TBQL) method [17] are employed for comparisons to illustrate the superiority of the dual ADP method. Table 1 presents the comparison results, which show that the dual ADP method can obtain better results.

Table 1 Comparison result.

Method	Cost (cent)	Saving rate (%)
Original	10,565.33	—
ASAGA	6923.64	34.47
MPC	6856.89	35.10
TBQL	6651.77	37.04
ADP	6585.49	37.67

V. Conclusion

This paper presents a novel smart grid system, which considers a variety of energy management approaches, and proposes a new self-learning dual ADP method to provide more economic benefits to users. The initial iterative control law is obtained by pre-training to accelerate the convergence of the algorithm and to improve the performance of the results. The results obtained by the dual ADP method are better than those obtained by other common methods through comparison experiments. In the future work, we will further analyze the performance of the method and consider the collaborative optimization problem between more smart grid systems.

Acknowledgment

This work was supported by the National Key R&D Program of China (No. 2021YFE0206100), the National Natural Science Foundation of China (No. 62073321), the National Defense Basic Scientific Research Program (No. JCKY2019203C029), and the Science and Technology Development Fund, Macao SAR (Nos. 0060/2021/A2 and 0015/2020/AMJ).

References

[1] M. Yousefi, A. Hajizadeh, M. N. Soltani, and B. Hredzak, Predictive home energy management system with photovoltaic array, heat pump, and plug-in electric vehicle, *IEEE Trans. Ind. Inf.*, 2021, 17(1), 430–440.

[2] S. Lee and D. H. Choi, Federated reinforcement learning for energy management of multiple smart homes with distributed energy resources, *IEEE Trans. Ind. Inf.*, 2022, 18(1), 488–497.

[3] L. Yu, W. W. Xie, D. Xie, Y. L. Zou, D. Y. Zhang, Z. X. Sun, L. H. Zhang, Y. Zhang, and T. Jiang, Deep reinforcement learning for smart home energy management, *IEEE Internet Things J.*, 2020, 7(4), 2751–2762.

[4] Q. Yang and H. Wang, Privacy-preserving transactive energy management for IoT-aided smart homes via blockchain, *IEEE Internet Things J.*, 2021, 8(14), 11463–11475.

[5] F. J. Luo, W. C. Kong, G. Ranzi, and Z. Y. Dong, Optimal home energy management system with demand charge tariff and appliance operational dependencies, *IEEE Trans. Smart Grid*, 2020, 11(1), 4–14.

[6] D. Fuselli, F. De Angelis, M. Boaro, S. Squartini, Q. L. Wei, D. R. Liu, and F. Piazza, Action dependent heuristic dynamic programming for home energy resource scheduling, *Int. J. Electr. Power Energy Syst.*, 2013, 48, 148–160.

[7] H. Shuai, J. K. Fang, X. M. Ai, Y. F. Tang, J. Y. Wen, and H. B. He, Stochastic optimization of economic dispatch for microgrid based on approximate dynamic programming, *IEEE Trans. Smart Grid*, 2019, 10(3), 2440–2452.

[8] P. Werbos, Advanced forecasting methods for global crisis warning and models of intelligence, in *Neural Networks for Control*. Cambridge, MA, USA: MIT Press, 1977, 25–38.

[9] P. J. Webros, A menu of designs for reinforcement learning over time, in *Neural Networks for Control*. Cambridge, MA, USA: MIT Press, 1990, 67–95.

[10] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf, Discrete-time nonlinear HJB solution using approximate dynamic programming: Convergence proof, *IEEE Trans. Syst., Man, Cybern., Part B (Cybern.)*, 2008, 38(4), 943–949.

[11] H. G. Zhang, Y. L. Liang, H. G. Su, and C. Liu, Event-driven guaranteed cost control design for nonlinear systems with actuator faults via reinforcement learning algorithm, *IEEE Trans. Syst., Man, Cybern.: Syst.*, 2020, 50(11), 4135–4150.

[12] Q. L. Wei, Z. H. Liao, R. Z. Song, P. J. Zhang, Z. Wang, and J. Xiao, Self-learning optimal control for ice-storage air conditioning systems via data-based adaptive dynamic programming, *IEEE Trans. Ind. Electron.*, 2021, 68(4), 3599–3608.

[13] T. Huang and D. R. Liu, A self-learning scheme for residential energy system control and management, *Neural Comput. Appl.*, 2013, 22(2), 259–269.

[14] Q. L. Wei, G. Shi, R. Z. Song, and Y. Liu, Adaptive dynamic programming-based optimal control scheme for energy storage systems with solar renewable energy, *IEEE Trans. Ind. Electron.*, 2017, 64(7), 5468–5478.

[15] L. Dong, Y. F. Tang, H. B. He, and C. Y. Sun, An event-triggered approach for load frequency control with supplementary ADP, *IEEE Trans. Power Syst.*, 2017, 32(1), 581–589.

[16] Q. L. Wei, Z. H. Liao, and G. Shi, Generalized actor-critic learning optimal control in smart home energy management, *IEEE Trans. Ind. Inf.*, 2021, 17(10), 6614–6623.

[17] M. Boaro, D. Fuselli, F. De Angelis, D. R. Liu, Q. L. Wei, and F. Piazza, Adaptive dynamic programming algorithm for renewable energy scheduling and battery management, *Cogn. Comput.*, 2013, 5(2), 264–277.

[18] M. Z. Zhang and Y. D. Gu, Optimization of ice-storage air conditioning system with ASAGA, in *Proc. 2014 IEEE Workshop on Advanced Research and Technology in Industry Applications*, Ottawa, Canada, 2014, 1042–1046.

[19] T. Wang, H. Kamath, and S. Willard, Control and optimization of grid-tied photovoltaic storage systems using model predictive control, *IEEE Trans. Smart Grid*, 2014, 5(2), 1010–1017.

[20] T. Y. Lee, Operating schedule of battery energy storage system in a time-of-use rate industrial user with wind turbine generators: A

multipass iteration particle swarm optimization approach, *IEEE Trans. Energy Convers*., 2007, 22(3), 774–782.

[21] T. Yau, L. N. Walker, H. L. Graham, A. Gupta, and R. Raithel, Effects of battery storage devices on power system dispatch, *IEEE Trans. Power Appar. Syst*., 1981, PAS-100(1), 375–383.

Qinglai Wei received the PhD degree from Institute of Automation, Chinese Academy of Sciences, China, in 2002. He is currently a researcher at State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, China, and a professor at Institute of Systems Engineering, Macau University of Science and Technology, China. He has hosted and participated in numerous national research projects, and published numerous academic papers in international journals and conferences. His research interests include intelligent control theory, adaptive dynamic programming, and smart home energy management.

Zehua Liao received the PhD degree from University of Chinese Academy of Sciences, China, in 2018. He is currently a researcher at State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, China, and at Institute of Systems Engineering, Macau University of Science and Technology, China. He has participated in several research projects and published several academic papers. His research interests include adaptive dynamic programming, intelligent optimization algorithms, and optimal management of smart grid systems.