

Received April 28, 2020, accepted May 19, 2020, date of publication June 8, 2020, date of current version June 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3000784

# Research on AGC Performance During Wind Power Ramping Based on Deep Reinforcement Learning

DONGYING ZHANG<sup>1</sup>, (Member, IEEE), HUITING ZHANG<sup>ID</sup><sup>1</sup>, (Student Member, IEEE),  
XU ZHANG<sup>ID</sup><sup>1</sup>, (Member, IEEE), XIAOYU LI<sup>1</sup>, KAIQI REN<sup>1</sup>, (Graduate Student Member, IEEE),  
YONGXU ZHANG<sup>1</sup>, AND YUN GUO<sup>2</sup>, (Student Member, IEEE)

<sup>1</sup>School of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China

<sup>2</sup>State Grid Rizhao Power Supply Company, Rizhao 276800, China

Corresponding author: Huiting Zhang (zht\_95@163.com)

This work was supported by the State Grid Corporation of China Project (Research and Application of Risk Assessment and Early Warning Technology for UHV Transmission On-Line Safety Operation of New Energy Base) under Grant 5215001600V4.

**ABSTRACT** With the increase in wind power penetration, wind power ramping events have increasingly influenced tie line power control in the power grid. Large power changes during ramping events make it difficult to accurately track the scheduling plans of tie lines and can even lead to overrun. Determining how to evaluate the control performance of automatic generation control (AGC) for wind power ramping has become an urgent problem. In this context, this paper studies the control performance of AGC for wind power ramping based on deep reinforcement learning. First, a tie line power control model of a power system with an AGC module is established. Then, measured data, which include thermal power, wind power, hydropower output and tie line power data, and a deep reinforcement learning method are combined for AGC parameter estimation based on the deep Q-network (DQN) algorithm. Next, the AGC parameter in different scenarios are fit by using measured phasor measurement unit (PMU) data, and on the basis of the fitted model, AGC performance evaluation is performed for wind power ramping events. Finally, the simulation results verify the feasibility and effectiveness of analysing the relationship between wind power ramping and AGC performance based on the AGC parameter fitting model.

**INDEX TERMS** Wind power ramping, tie line, AGC, measured data, deep reinforcement learning.

## I. INTRODUCTION

In recent years, renewable energy resources such as wind power have begun to connect to the power grid on a large scale. However, due to the randomness and uncertainty of wind power, the stability of power system operations will be tested as wind power penetration increases [1]. Notably, the occurrence of wind ramp events [2]–[6] will adversely affect the power balance of the system, resulting in the tie lines of interconnected power systems having difficulty in accurately tracking the scheduling plan, thus increasing the difficulty of tie line power control. In this context, it must be determined whether automatic generation control (AGC) [7]–[9] can reduce the impact of wind ramp events on the power of tie lines, and evaluations of the

control performance of AGC for wind power ramping are important.

The problems related to AGC can be solved by digital simulation technology. In terms of AGC simulation model research, reference [10] successively introduced an AGC model built with the power system simulation programs PSS/E, Eurostag and PSD-FDS. Reference [11] proposed a dynamic AGC simulation method based on a smart grid for AGC performance quantification. Reference [12] used a hybrid system modelling method to establish an AGC model that can be applied to the entire dynamic process simulation of a power system. However, the parameters of the AGC models in the above references are all fixed; therefore, these methods can only be used in the simulation of simple scenes. When AGC is implemented, complex nonlinear characteristics are observed at long time scales, and the above modelling methods cannot effectively reflect these characteristics.

The associate editor coordinating the review of this manuscript and approving it for publication was Manoj Datta<sup>ID</sup>.

AGC simulation models based on data-driven methods can encompass the complex characteristics of AGC. Such an approach is suitable for the simulation of complex scenes and can effectively evaluate the control performance of AGC for ramping events. However, the effect of AGC based on data-driven methods must be evaluated; that is, the AGC parameters must be fit so that the relationships among the model output variables are consistent with those in an actual situation. At present, large-scale plants and stations are generally equipped with phasor measurement units (PMU) [13], which can obtain real-time operation data for a power system. Therefore, the AGC parameters can be fitted by using the measured PMU data, and the control effect of AGC for tie lines under different wind power penetration rates and ramp rates can be further evaluated. AGC parameter fitting is essentially a decision-making problem, and deep reinforcement learning [14]–[16] can be used to analyse data characteristics and make decisions to solve this problem. Deep reinforcement learning combines the powerful perceptual comprehension ability of deep learning and decision-making ability of reinforcement learning to achieve one-to-one correspondence from perception to action [17]–[19]. Deep learning uses information from the environment to extract features and generate state representations of the current environment. Reinforcement learning selects actions based on the current state to achieve the expected goal. At present, deep reinforcement learning has yielded remarkable achievements related to many aspects of power systems, such as the coordinated control of the hybrid energy storage in microgrids [20], the generation unit tripping strategy under emergency circumstances [21] and AGC strategy research [22], thereby displaying its effectiveness in solving decision-making problems.

This paper proposes a data-driven method for AGC parameter fitting based on deep reinforcement learning to evaluate the control performance of AGC for wind power ramping. First, by analysing the operation data from a regional power grid, a data-driven power grid model with AGC is established. Second, based on the deep Q-network (DQN) algorithm, the model framework for AGC performance simulation is established, and the AGC parameters are established. Then, the AGC performance under different wind power penetration rates and ramp rates is evaluated based on measured data. Finally, the simulation study shows the effectiveness and accuracy of the proposed method.

## II. AGC MODEL BASED ON DATA DRIVE

### A. DATA CHARACTERISTICS

This paper analyses thermal power output, hydropower output, wind power output and tie line power data from a certain area after AGC is put into operation. As shown in Figure 1, the wind power output in this area, which accounts for a relatively high proportion of the power output, is characterized by strong random fluctuations, and some wind ramping events occur. The proportion of hydropower output is low,

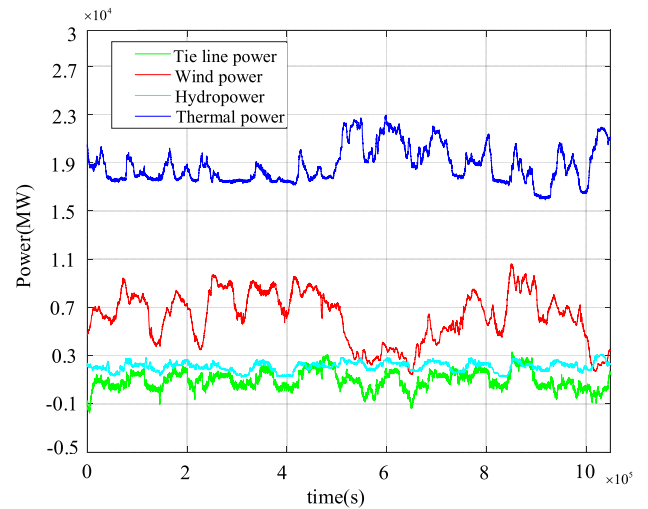


FIGURE 1. Output of various types of units and tie line power.

and the output change is stable. When the wind power output suddenly changes, the trend of the thermal power output is approximately opposite to that of wind power. It can be seen that under the control of AGC, the output power of the thermal power unit is mainly adjusted to balance the active power of the system, and thus, the tie line power maintains a relatively stable state.

In the above figure, due to the large amount of data and the different magnitudes of each group of data, it is impossible to analyse the relationship between variables accurately from the graph. Therefore, in order to further understand the regulation of tie lines with AGC, the relationship between the wind power output and tie line power is analysed by using a scatter density diagram, as shown in Figure 2. The trend reflected by scatter is the relationship between wind power output and tie line power, and the colour depth reflects the data density to observe the data distribution characteristics. In Figure 2, the highlighted part of the figure represents the

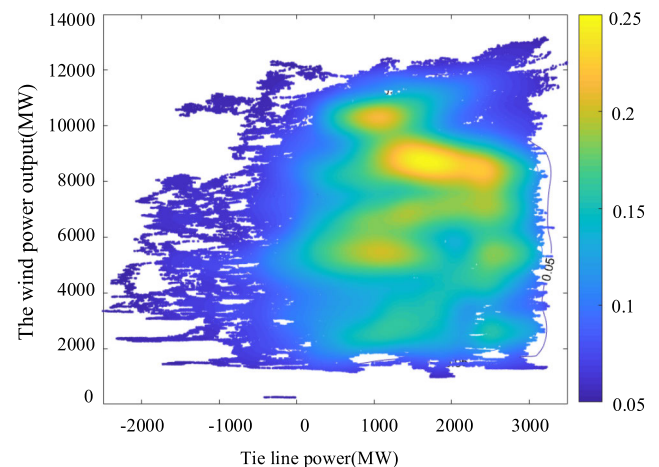
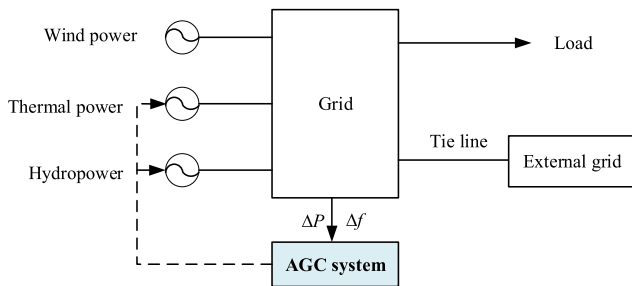


FIGURE 2. Relationship between wind power output and tie line power.

concentrated distribution area of data, and 0.05-0.25 on the colour bar represents the data density. The distributions of the wind power output and tie line power are not uniform. The tie line power is mostly distributed between 160 MW and 3000 MW, and the wind power output is concentrated between 2000 MW and 15000 MW. In this range, there are many highlighted areas, and the relationship between wind power output and tie line power in different areas is different. Overall, the relationship between the two is complex and has strong nonlinear characteristics. It is difficult to accurately use a mathematical model to express the regulation effect of AGC on the tie line when wind power output changes in multiple scenarios, and this needs to be analysed in combination with the nonlinear characteristic fitting ability of deep learning.

**B. DATA-DRIVEN AGC MODELLING**

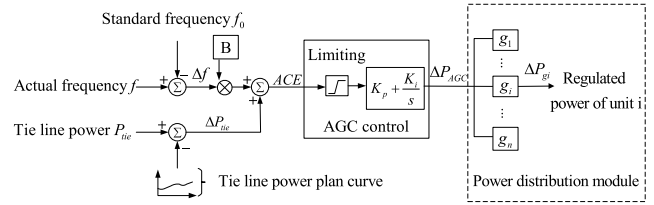
Data-driven AGC modelling should reflect the mapping relationship for each group of data, that is, the relationships among the thermal power, hydropower, wind power output, and tie line power. The key step in AGC modelling is simplifying the internal structure of the power grid, with a focus on the output of each unit and the tie line power changes in the AGC model. In this paper, the internal structure of the power grid is regarded as a “black box”, and the AGC model, various types of units, grid load and tie lines are retained. The basic framework of the power system model with AGC is shown in Figure 3. In Figure 3, the “Grid” is the regional power grid that provides wind power output, hydropower output, thermal power output and tie line power. The “External grid” is the other regional power grid that has power exchange with the “Grid”. The “Grid”, “Tie line” and “External grid” together constitute the two regional interconnection power systems.



**FIGURE 3. Framework diagram of the power system model with AGC.**

AGC takes the frequency deviation and tie line power deviation as the inputs, and outputs unit regulation power through the controller to balance the active power of the system, thereby reducing the frequency deviation and tie line power deviation. A typical AGC model is shown in Figure 4.

In Figure 4,  $\Delta f$  is the frequency deviation within the region,  $\Delta P_{tie}$  is tie line power deviation,  $B$  is the frequency deviation coefficient, and  $g_i$  is the distribution coefficient

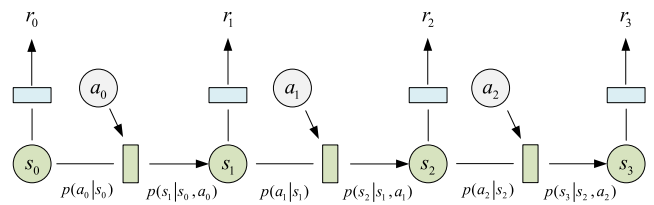


**FIGURE 4. AGC model.**

of the  $i$ -th unit. The AGC controller is a PI controller, and  $K_p$  and  $K_i$  are the proportionality coefficient and integral coefficient, respectively. The input signal of the controller is the area control error (ACE), which includes the area frequency deviation and tie line exchange power deviation.

$$ACE = B\Delta f + \Delta P_{tie} \tag{1}$$

The basic objective of AGC is to control tie line power by inputting the ACE into the PI controller and outputting the power setting value  $\Delta P_{AGC}$ . Then, the power distribution module assigns  $\Delta P_{AGC}$  to each AGC unit. The unit will adjust the power according to the instructions issued by the AGC for the power control of tie lines. Therefore, the parameters of the AGC controller and power distribution module play an important role in the tie line power control of power grids and affect the control performance of AGC. To develop an accurate AGC model, it is necessary to fit parameters such as the proportionality coefficient  $K_p$ , integral coefficient  $K_i$  and unit regulation power distribution coefficient  $g$  with deep reinforcement learning.



**FIGURE 5. Markov decision process.**

**III. DEEP REINFORCEMENT LEARNING**  
**A. REINFORCEMENT LEARNING**

Reinforcement learning is a process of interaction between agents and their environment. Agents continuously try to make mistakes by executing various behaviours that yield the maximum cumulative reward. This learning process can be transformed into a Markov decision process (MDP). A MDP can be represented by four tuples  $(s, a, p, r)$  composed of state  $s$ , action  $a$ , state transition probability  $p$ , and return  $r$ , and the corresponding decision-making process is shown in Figure 5. At the end of the process, the agent will receive an action sequence, called a strategy, which will be recorded. Then, the cumulative reward of the strategy will be returned,

as shown in Equation (2).

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \gamma^3 r_{t+4} + \dots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \quad (2)$$

In Equation (2),  $\gamma$  is the discount factor, indicating the weight of future awards in the cumulative award.

To maximize the reward value of the strategy  $\pi$ , it is necessary to evaluate the value of action  $a$  under a specific state  $s$ . The state action value function  $Q_{\pi}(s, a)$  can be used to evaluate the action behaviour:

$$Q_{\pi}(s, a) = E_{\pi}[R_t | s_t = s, a_t = a] \quad (3)$$

In this case,  $Q_{\pi}(s, a)$ , or Behrman's equation, can be expressed as follows:

$$Q_{\pi}(s, a) = E_{\pi}[r_{t+1} + \gamma Q_{\pi}(s_{t+1}, a_{t+1}) | s_t, a_t] \quad (4)$$

The above formula indicates that the state action value function can be obtained by continuous iterative calculations, and thus, the iterative Bellman equation can be used to solve the Q value function. The strategy with the largest value function is called the optimal strategy.

### B. DQN ALGORITHM

When the state space is large, it is unrealistic to use the iterative method to solve for the optimal strategy. At this time, it is necessary to use a function approximation method to express the value function. The DQN algorithm uses a deep neural network with weight  $\theta$  as an approximate representation of the current value function. Then, the loss function according to the correct Q value provided by reinforcement learning is obtained. By minimizing the loss function, the network weight  $\theta$  is continuously updated. The loss function  $L(\theta)$  and the weight  $\theta$  are updated as follows:

$$L(\theta) = E[(Y_i - Q(s_t, a_t; \theta))^2] \\ = E[(r_{t+1} + \gamma \max_a Q(s_{t+1}, a_{t+1}; \theta) - Q(s_t, a_t; \theta))^2] \quad (5)$$

$$\theta_{t+1} = \theta_t + \alpha(r_{t+1} + \gamma \max_a Q(s_{t+1}, a; \theta) - Q(s_t, a_t; \theta)) \nabla_{\theta} Q(s_t, a_t; \theta) \quad (6)$$

where  $Y_i$  is the optimization objective of the value function, which is the objective Q value;  $Q(s_t, a_t; \theta)$  is the evaluation of  $Q(s_t, a_t)$ ; and  $\alpha$  is the learning rate. To reduce the correlation between the current network Q value and the target  $Y_i$  and improve the stability of the algorithm, the DQN algorithm uses a separate network, that is, a target network, to generate the target Q value in the training process. During the training process, the parameters of the current network  $\theta$  are updated in real time, and the parameters of the target network  $\theta'$  are unchanged. Only after the c-step iteration are the parameters of the current network copied to the target network, and the loss function is:

$$L_i(\theta_i) = E[r_{t+1} + \gamma \max_a Q(s_{t+1}, a_{t+1}; \theta'_i) - Q(s_t, a_t; \theta_i)]^2 \quad (7)$$

where  $\theta_i$  is the parameter of the current network at the i-th iteration and  $\theta'_i$  is the parameter of the target network.

The current network parameters  $\theta$  and target network parameters  $\theta'$  by random gradient descent are updated as follows:

$$\nabla_{\theta} L(\theta) = E(r + \gamma \max_{a'} Q(s_{t+1}, a_{t+1}; \theta') - Q(s_t, a_t; \theta)) \nabla_{\theta} Q(s_t, a_t; \theta) \quad (8)$$

$$\theta_{i+1} = \theta_i + \alpha \nabla_{\theta} L_i(\theta_i) \quad (9)$$

$$\theta'_i = \theta_{i+c} \quad (10)$$

In addition, in the DQN algorithm, agents typically use  $\epsilon$ -greedy strategies to select actions in each step; that is, actions are randomly selected under the probability of  $\epsilon$ , and the actions with the highest Q values in the current state are selected at the probability of  $1 - \epsilon$ , as shown in the following formula:

$$\pi(a | s) = \begin{cases} \frac{\epsilon}{m} + 1 - \epsilon & \text{if } a = \arg \max Q(s, a) \\ \frac{\epsilon}{m} & \text{others} \end{cases} \quad (11)$$

where m is the total number of all optional actions.

### IV. AGC PARAMETER FITTING BASED ON THE DQN ALGORITHM

This paper uses the DQN algorithm in deep reinforcement learning to fit the AGC parameters. The fitting process can be represented by the reinforcement learning process, as shown in Figure 6. The AGC simulation model and measured data are used as the environment, the DQN algorithm is used as the agent, and parameter fitting is achieved based on the interactions between the agent and environment. In this process, the agent obtains the current state of the power grid from the environment, determines the AGC parameter values according to a certain strategy, and acts on the AGC module to obtain the next power grid state. Then, according to the comparison between the state and the measured data,

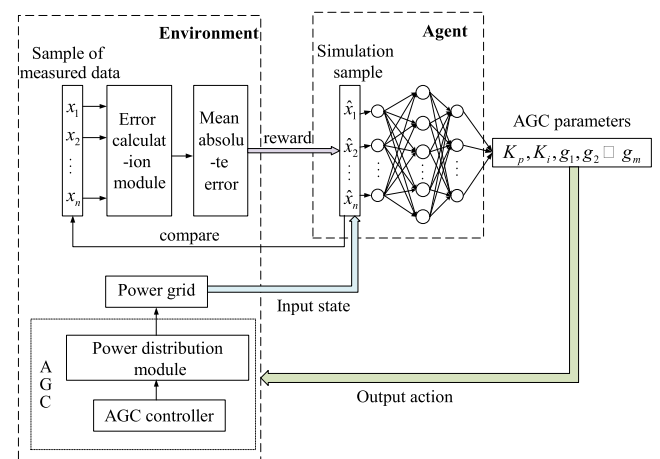


FIGURE 6. AGC parameter fitting framework.



the environment will return rewards to the agent. This process is repeated until the desired goal is achieved.

### A. DESIGN OF THE STATE SPACE, ACTION SPACE AND REWARD FUNCTION

The state space is composed of the outputs of various types of units in the power system model at each time step, including the thermal power output, hydropower output, wind power output and tie line power, namely:

$$S_t = \{P_{fire}^t, P_{water}^t, P_{wind}^t, P_{tie}^t\} \quad (12)$$

The action space parameters include the following AGC model parameters: the proportionality coefficient, integral coefficient and power distribution coefficient of each unit; these parameters can be expressed as follows:

$$A = [a_1 \ a_2 \ \dots \ a_{m+2}] = [K_p, K_i, g_1, g_2, \dots, g_m] \quad (13)$$

where  $g$  is the power distribution coefficient of the unit and  $m$  is the number of units.

The reward function involves reward and punishment processing based on whether the output power of the model after an action is close to the measured value. By comparing the output of each unit and the tie line power with the actual data, we can calculate the average error and use it as the basis for evaluating the action strategy. The average error is calculated as follows:

$$error1 = \frac{1}{n_{fire}} \sum_1^n |P_{fire} - P_{fire-true}| \quad (14)$$

$$error2 = \frac{1}{n_{water}} \sum_1^n |P_{water} - P_{water-true}| \quad (15)$$

$$error3 = \frac{1}{n_{wind}} \sum_1^n |P_{wind} - P_{wind-true}| \quad (16)$$

$$error4 = \frac{1}{n_{line}} \sum_1^n |P_{line} - P_{line-true}| \quad (17)$$

In the formula,  $P_{fire}$ ,  $P_{water}$ ,  $P_{wind}$  and  $P_{line}$  are the outputs of the thermal power unit, hydropower unit, wind power unit and tie lines, respectively.  $P_{fire-true}$ ,  $P_{water-true}$ ,  $P_{wind-true}$  and  $P_{line-true}$  are the actual outputs and tie line power of the corresponding units. According to the error obtained from Equations (14)-(17), set the reward function as follows:

$$r(s_t, a_t) = \begin{cases} -1; & error1_t \text{ and } error2_t \text{ and } error3_t \text{ and } error4_t > 0.2 \\ 1; & 0.04 < error1_t \text{ and } error2_t \text{ and } error3_t \text{ and } error4_t < 0.1 \\ 2; & error1_t \text{ and } error2_t \text{ and } error3_t \text{ and } error4_t < 0.04 \\ 0; & others \end{cases} \quad (18)$$

### B. LEARNING PROCESS

In the DQN algorithm, both the current network and the target network use a deep convolution neural network structure. The two networks have the same structure, and only the parameters are different. To ensure the integrity of data, the pooling process is omitted in the convolutional neural network, and only two convolutional layers are used for feature extraction. Among them, the first layer selects  $(4 \times 2 \times 1)$  convolution kernels, and the second layer selects  $(8 \times 4 \times 1)$  convolution kernels. The fully connected layer consists of two hidden layers, in which the numbers of hidden neurons are 30 and 60. The activation function is the ReLU (rectified linear unit) function, and the number of neurons in the output layer is the number of action values needed to output Q under different actions.

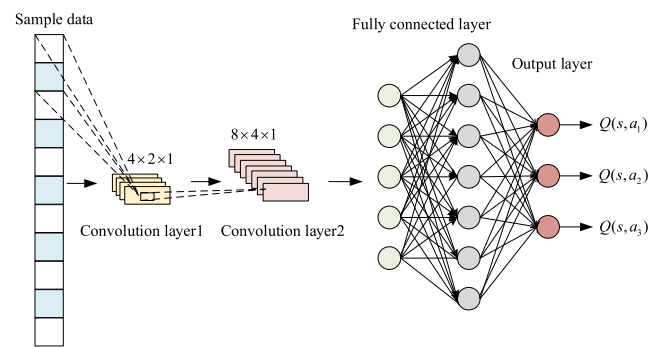


FIGURE 7. Deep convolutional neural network model.

The reinforcement learning process aims to maximize the expected rewards, and through continuous learning, we can obtain the maximum reward. During the learning process, the output power of the T-time grid model  $s_t$  is first input into the deep convolution neural network model. After two convolutional and fully connected layer steps, the output layer of the neural network outputs the Q value of each AGC parameter. Then, according to the strategy shown in Equation (11), the agent selects the AGC parameters  $a_{t+1}$  that need to be changed, transmits them to the AGC model to obtain the next power output state of the grid  $s_{t+1}$ , and determines the reward after taking the action according to the reward function. Finally, the loss function  $L(\theta)$  is calculated with Equation (7), and the parameters of the current network and target network are updated by applying Equations (8)-(10).

### C. PARAMETER SETTING

As shown in Figure 6, the DQN algorithm model needs to match the AGC parameters in combination with the power grid model. Therefore, the initial parameter values should be set for the AGC model, as shown in Table 1. These values are randomly set within the setting range of actual unit parameters and AGC parameters.

To make the training process of the algorithm stable and accelerate the learning and convergence speeds of the algorithm, through simulation verification, the values of

TABLE 1. AGC initial operation parameters.

Symbol	Parameter	Value
Kp	proportionality coefficient	5
Ki	integral coefficient	1
g1	G1 power distribution coefficient	0.26
g2	G2 power distribution coefficient	0.21
g3	G3 power distribution coefficient	0.23
g4	G4 power distribution coefficient	0
g5	G5 power distribution coefficient	0
g6	G6 power distribution coefficient	0.22
g7	G7 power distribution coefficient	0.08
g8	G8 power distribution coefficient	0

TABLE 2. DQN algorithm parameter setting.

Symbol	Parameter	Value
$\alpha$	learning rate	0.5
$\gamma$	discount factor	0.9
$\epsilon$	strategy selection probability	0.9
memory_size	memory capacity	1000
replace_target_iter	target network replacement steps	50

each parameter in the DQN algorithm are set as shown in Table 2.

D. ALGORITHM FLOW

The AGC simulation flow chart based on the DQN algorithm is shown in Figure 8.

V. EXAMPLE ANALYSIS

A. SIMULATION MODEL

Referring to the AGC model framework shown in Figure 4, this paper builds the power system model with AGC shown in the following figure based on the PowerFactory/DIGSILENT simulation platform to obtain operation data using different AGC parameters. The rated active power value of each generator in the model is set according to the actual unit nameplate parameters. See Table 3 for the specific unit information.

B. SAMPLE DATA

In this paper, the wind power, hydropower, thermal power output and tie line power recorded in seconds of a certain region in May 2019 are used as sample data to analyse tie line power fluctuations in the case of wind power ramping up, ramping down and no ramping.

According to [2], during the time period  $(t, t + \Delta t)$ , when the absolute value of the difference between wind power at the first and last moments exceeds a given threshold, the wind power ramping event occurs. If the power increases, it is a ramp-up event, and if the power decreases, it is a ramp-down event. In general,  $\Delta t$  is taken as 30 minutes, and the threshold

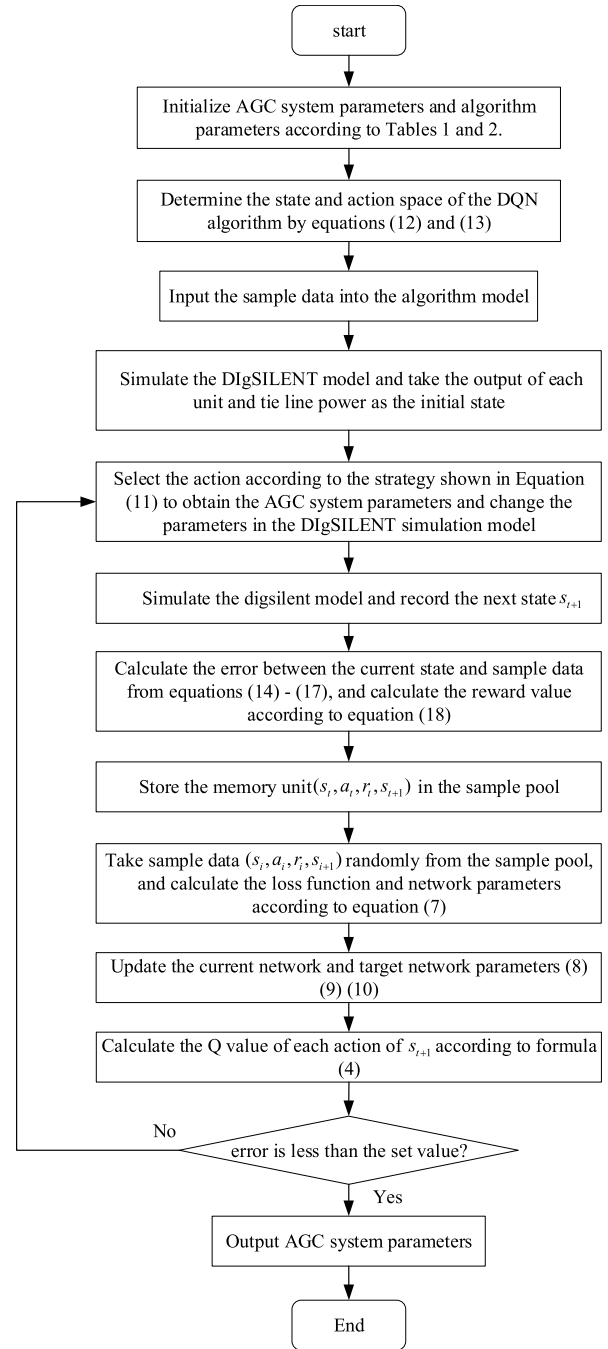


FIGURE 8. Algorithm flowchart.

values of ramp-up events and ramp-down events are 20% and 15% of the installed capacity of wind power respectively. However, the control cycle of the AGC system is generally 6-10 minutes, and thus, this paper selects 10 minutes as the control cycle length, 7% and 5% of the installed capacity of wind power as the threshold values, and slices the data containing the ramping events in 10-minute intervals to obtain one-way continuous wind power ramping events. At last, three conditions are obtained: ramping up, ramping down and no ramping of wind power. See Table 4 for details.

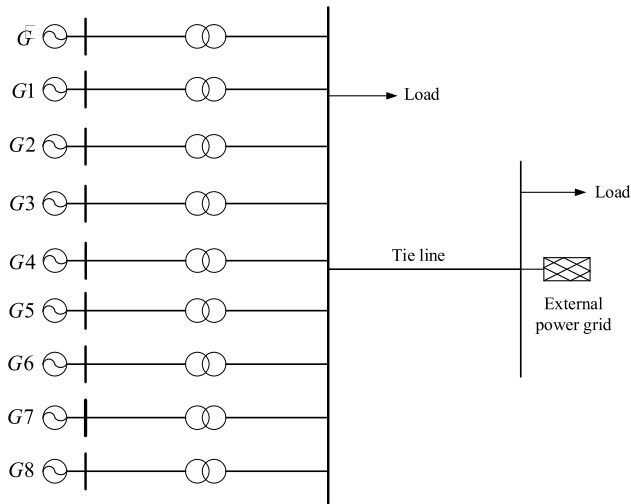


FIGURE 9. Power system simulation model.

TABLE 3. Generator parameters.

Symbol	Type of generator	Rated active power (MW)	Number of parallel units
$G_0$	wind power	2.2	5000
$G_1$	thermal power	1100	8
$G_2$	thermal power	660	13
$G_3$	thermal power	150	34
$G_4$	thermal power	350	13
$G_5$	hydropower	21	25
$G_6$	hydropower	34	14
$G_7$	hydropower	94	15
$G_8$	hydropower	15	55

TABLE 4. Three working conditions.

Working condition	Wind power ramping situation	Number of slices	Percentage
Working condition 1	wind power ramping down	459	19.13%
Working condition 2	wind power ramping up	715	29.79%
Working condition 3	no ramping	1226	51.08%

It can be seen from the above table that the probability of occurrence of ramp-up events is greater than that of ramp-down events. When the wind power ramping amplitude reaches a certain critical value, the control performance of the AGC on the tie line will be affected. To fit the control performance of AGC under wind power ramping, the data for condition 1 and condition 2 are selected as the sample data input into the algorithm. The relationship between the output of each unit and tie line power for condition 1 and condition 2 is shown in Figure 10 and Figure 11, respectively.

In Figure 10 and Figure 11, the output data for the tie line, wind power, hydropower and thermal power are normalized

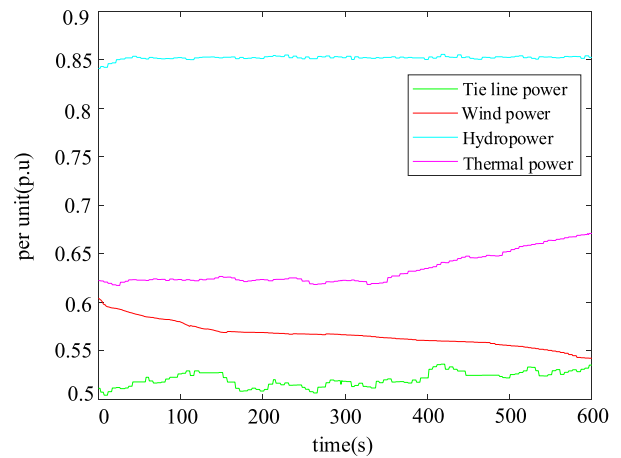


FIGURE 10. Working condition 1.

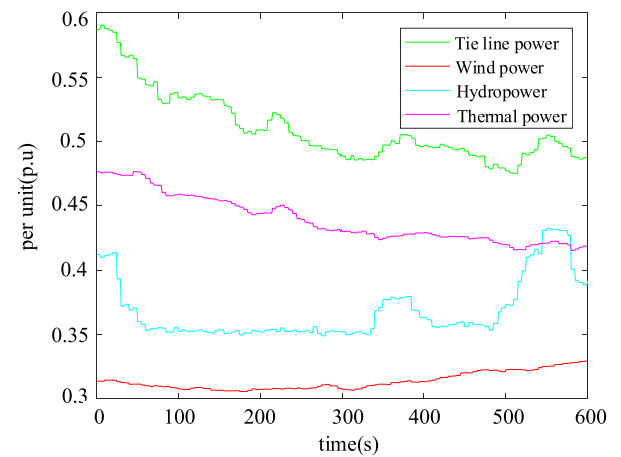


FIGURE 11. Working condition 2.

by the method of deviation standardization (min-max); that is, the output is mapped to the range of [0,1] based on the minimum value of each data set, and the curve in the figure reflects the wave motion of each output. The two figures show that when the wind power ramps up or down, the thermal power output increases or decreases accordingly, which makes the tie line power relatively stable.

### C. ANALYSIS OF RESULTS

According to the algorithm flow chart shown in Figure 8, the sample data are input into the algorithm. The training model includes 1000 episodes, and the training of an episode includes multiple training steps. Each step returns a reward value, which is summed to obtain the overall value for an episode. Figure 12 and Figure 13 show the number of steps and the average reward value of each episode in the training process. The number of training steps in the first 50 episodes is large, and the number of training steps after 100 episodes is greatly reduced and tends to be stable. Correspondingly, the average reward value initially increases rapidly and then slowly increases. After approximately 200 episodes,

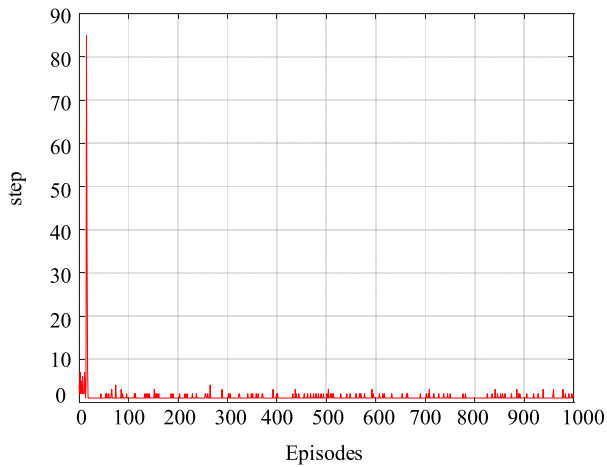


FIGURE 12. DQN algorithm training steps.

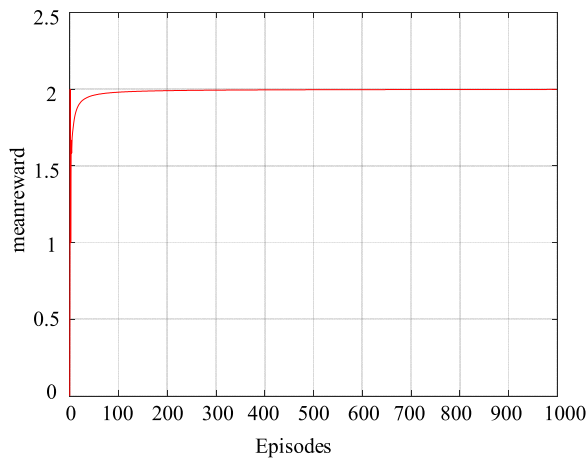


FIGURE 13. DQN algorithm average reward value.

the reward value is stable at the maximum expected value, with good convergence ability.

After the AGC model is fully trained with the DQN algorithm, the action strategy with the maximum return value is obtained. Based on this strategy, the model is run to obtain the real-time unit output and tie line power. To verify the fitting effect of the DQN algorithm on AGC parameters, the particle swarm optimization algorithm (PSO) is introduced for test comparison. The comparison of the wind power output, tie line power and actual data of the working condition 1 with different algorithms is shown in Figure 14 and Figure 15. It can be seen that the wind power output and tie line power using the DQN algorithm are closer to the actual data, indicating that the fitting effect of the AGC parameters fitted by the DQN algorithm is similar to the actual system, which verifies the effectiveness of the proposed method.

#### D. AGC PERFORMANCE EVALUATION

In most current simulation models, the setting of AGC parameters is not based on the actual situation. Therefore, the AGC

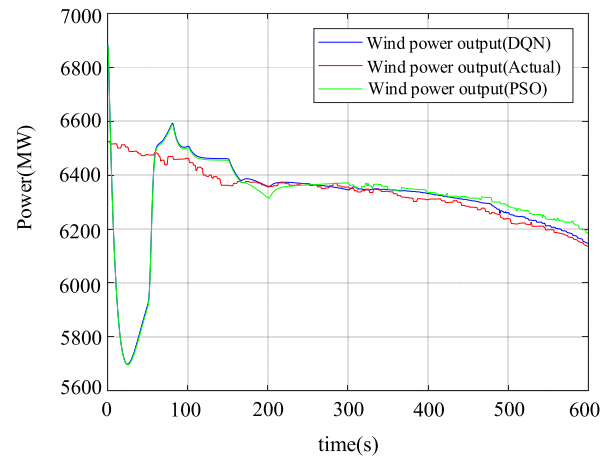


FIGURE 14. Comparison between the wind power output of DQN and the PSO algorithm and the actual output.

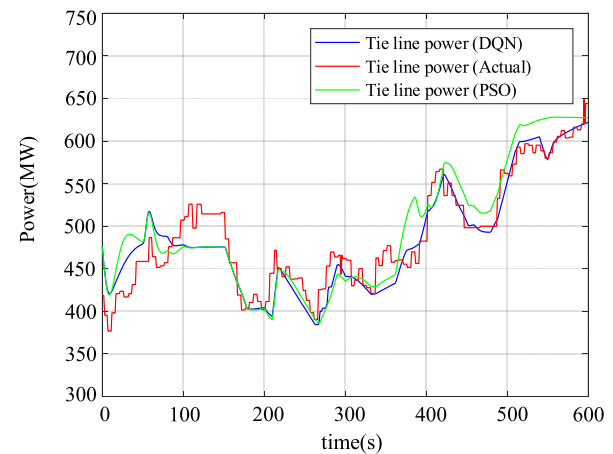
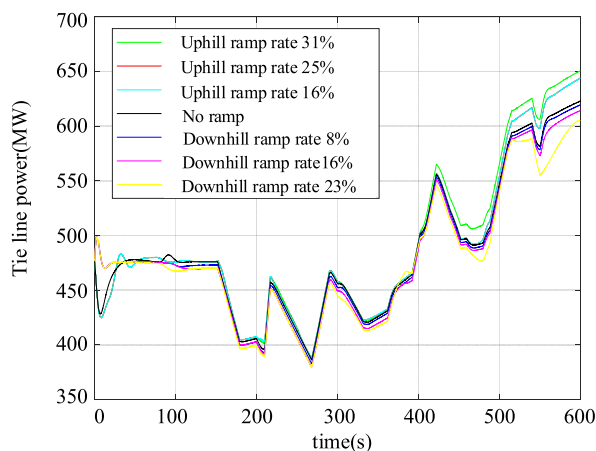


FIGURE 15. Comparison between the tie line power of DQN and the PSO algorithm and the actual tie line power.

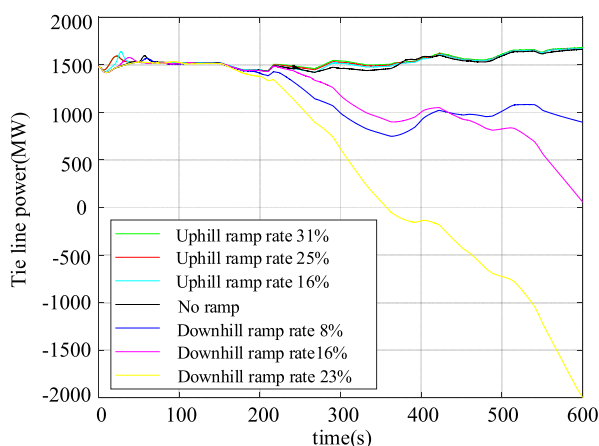
action behaviour is not consistent with the actual situation and cannot accurately reflect the process of secondary frequency modulation, which in turn affects the accuracy of the scene analysis in which AGC participates. In the above, we used the actual data to fit the parameters that reflect the actual AGC adjustment characteristics through the DQN algorithm. On this basis, we can analyse the actual AGC adjustment performance in various scenarios, such as wind power ramping. Wind power ramping events have a considerable impact on AGC performance. To evaluate the AGC performance after parameter fitting, this paper explores wind power ramping events with different wind power penetration rates based on the power system model shown in Figure 9 and assesses the corresponding regulation ability according to the changes in the tie line power. Figure 16 and Figure 17 show the changes in the tie line power at different ramp rates for 20% and 50% wind power penetration, respectively.

In Figure 16 and Figure 17, when the wind power penetration rate is 20%, the tie line power at different ramp





**FIGURE 16.** Tie line power at different ramp rates with 20% wind power penetration.



**FIGURE 17.** Tie line power at different ramp rates with 50% wind power penetration.

rates does not change significantly compared to that without ramping, and the possibility of deviation from the planned value is small. When the penetration level reaches 50% and a ramp-down event occurs, the tie line power will fluctuate, and the possibility of deviation from the planned value will increase compared with that for no-ramping and ramp-up events. Therefore, with the increase in penetration, the probability of the tie line power deviating from the planned value increases.

Figure 17 shows that when the wind power penetration level is high, the impact of a ramp-down event on the tie line power is much greater than that for a ramp-up event. The larger the ramp-down rate is, the larger the fluctuation range of the tie line power, and reverse power transmission can occur. In this case, the risk of the tie line power not reaching the planned value sharply increases.

## VI. CONCLUSION

In this paper, a data-driven concept is used to apply deep reinforcement learning to AGC parameter fitting. A data-driven

AGC parameter fitting method based on deep reinforcement learning is proposed. The main work can be summarized as follows:

- 1) A power control model of tie lines with an AGC module is established. Based on the model, the relationship between wind power ramping and the tie line is analysed.
- 2) In this paper, deep reinforcement learning is introduced into AGC parameter fitting, and the DQN algorithm is used to fit the AGC parameters in various scenarios combined with the measured data of the power grid.
- 3) Based on the model of the measured data fitting and data-driven tie line, the control performance of AGC for wind power ramp events is evaluated, and the effectiveness of the proposed method is verified.

The research focus of this paper is AGC parameter fitting, which is the basis for evaluating the control performance of AGC. From this approach, we can assess the influence of wind power ramping on the control of tie lines, which plays an important role in preventing tie line limits from being exceeded and in developing new control strategies.

## REFERENCES

- [1] Y. S. X. Xue Lei and F. Xue, "A review on impacts of wind power uncertainties on power systems," *Proc. CSEE*, vol. 34, no. 29, pp. 5029–5040, Oct. 2014.
- [2] R. Sevlian and R. Rajagopal, "Detection and statistics of wind power ramps," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3610–3620, Nov. 2013.
- [3] J. Freedman, M. Markus, and R. Penc, *Analysis of West Texas Wind Plant Ramp-Up and Ramp-Down Events*. New York, NY, USA: Albany, 2008.
- [4] H. Ma, C. H. Li, and Y. T. Liu, "Multi-level interval early warning method for wind power ramping event," *Autom. Electr. Power Syst.*, vol. 41, no. 11, pp. 39–47, Jun. 2017.
- [5] J. Zhang, C. Lu, J. Song, and J. Zhang, "Real-time AGC dispatch units considering wind power and ramping capacity of thermal units," *J. Modern Power Syst. Clean Energy*, vol. 3, no. 3, pp. 353–360, Jul. 2015.
- [6] H. Ma and Y. T. Liu, "IGDT robust model-based coordinated scheduling strategy for wind power ramp events," *Proc. CSEE*, vol. 36, no. 17, pp. 4580–4589, Sep. 2016.
- [7] W. Z. Zhao, B. Li, H. Wei, C. F. Wei, and J. Deng, "The optimal AGC control strategy considering the primary frequency regulation under the control performance standard for the interconnected power grid," *Proc. CSEE*, vol. 36, no. 10, pp. 2656–2664, May 2016.
- [8] Z. H. Gao, G. Chen, J. F. Yang, X. L. Teng, and J. G. Yao, "Power control of tie line in UHV interconnected power grid (1) AGC control strategy," *Autom. Electr. Power Syst.*, vol. 33, no. 15, pp. 51–55, Aug. 2009.
- [9] Z. H. Gao, G. Chen, X. L. Teng, X. B. Zhang, and J. G. Yao, "Power control of tie line in UHV interconnected power grid (1) AGC performance evaluation," *Autom. Electr. Power Syst.*, vol. 33, no. 16, pp. 61–64, Aug. 2009.
- [10] H. Huo, W. Z. Zhong, Y. B. Chen, and J. Ma, "Research survey on control, modeling, simulation and application of automatic generation control system," *Shaanxi Electric Power*, vol. 42, no. 9, pp. 62–67, 2009.
- [11] L. Wang and D. Chen, "Extended term dynamic simulation for AGC with smart grids," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2011, pp. 1–7.
- [12] C. W. Xinli Song and W. Zhong, "Modeling of automatic generation control for power system transient medium-term and long-term stabilities simulations," *Power Syst. Technol.*, vol. 37, no. 12, pp. 3439–3444, Dec. 2017.
- [13] H. Sun, M. M. Wang, and X. Qi, "Synchronous generator parameters identification based on PMU data," *Power Syst. Protection Control*, vol. 42, no. 03, pp. 31–36, Feb. 2014.
- [14] K. Arulkumar, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "A brief survey of deep reinforcement learning," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 26–38, Nov. 2017.

- [15] Q. Liu, J. W. Zhai, Z. C. Zhang, S. Zhong, and J. Xu, "A survey on deep reinforcement learning," *Chin. J. Comput.*, vol. 41, no. 1, pp. 1–27, Jan. 2018.
- [16] J. Hou, H. Li, J. Hu, C. Zhao, Y. Guo, S. Li, and Q. Pan, "A review of the applications and hotspots of reinforcement learning," in *Proc. IEEE Int. Conf. Unmanned Syst. (ICUS)*, Oct. 2017, pp. 506–511.
- [17] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016.
- [18] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," 2015, *arXiv:1509.02971*. [Online]. Available: <https://arxiv.org/abs/1509.02971>
- [19] S. Hansen, "Using deep Q-learning to control optimization hyperparameters," Feb. 2016, *arXiv:1602.04062*. [Online]. Available: <https://arxiv.org/abs/1602.04062>
- [20] Z. D. Zhang, C. M. Qiu, D. X. Zhang, S. W. Xu, and X. He, "A coordinated control method for hybrid energy storage system in microgrid based on deep reinforcement learning," *Power Syst. Protection Control*, vol. 43, no. 6, pp. 1914–1921, Jun. 2019.
- [21] W. Liu, D. Zhang, X. Wang, J. X. Hou, and L. P. Liu, "A decision making strategy for generating unit tripping under emergency circumstances," *Proc. CSEE*, vol. 38, no. 1, pp. 109–119, Jan. 2018.
- [22] L. Xi, X. I. Lei, Y. U. Lu, and F. U. Yimu, "Automatic generation control based on deep reinforcement learning with exploration awareness," *Proc. CSEE*, vol. 39, no. 14, pp. 4150–4162, Jul. 2019.



**XU ZHANG** (Member, IEEE) received the M.S. and Ph.D. degrees in power systems and automation from North China Electric Power University, Beijing, China, in 2005 and 2013, respectively. He has been an Assistant Professor with the School of Electrical and Electronic Engineering, since 2005. His current research interests are in the areas of power system intelligent analysis and control, power system frequency regulation by wind power, and power grid fault diagnosis and restoration.



**XIAOYU LI** received the B.S. degree from North China Electric Power University, Beijing, China, in 2019, where he is currently pursuing the M.S. degree in power systems and automation. His research interests include new energy generation and grid connection.



**KAIQI REN** (Graduate Student Member, IEEE) received the B.S. degree from Shanxi University, Shanxi, China, in 2019. He is currently pursuing the M.S. degree in electrical engineering with North China Electric Power University, Beijing, China. His research field is the simulation of the equivalent inertia time constant of power systems.



**YONGXU ZHANG** received the B.S. degree from North China Electric Power University–Baoding, Baoding, China, in 2019. He is currently pursuing the M.S. degree in electrical engineering with North China Electric Power University, Beijing, China. His research interests include power system fault diagnosis and simulation analysis.



**YUN GUO** (Student Member, IEEE) received the B.S. degree from Northeast Electric Power University, Jilin, China, in 2017, and the M.S. degree in electrical engineering from North China Electric Power University, Beijing, China. She is currently working with State Grid Rizhao Power Supply Company, Rizhao, China. Her research interests are in the area of power system fault diagnosis based on deep learning.

...



power system analysis, and control and grid connected technology of new energy generation.

**DONGYING ZHANG** (Member, IEEE) received the Ph.D. degree in power systems and automation from the Harbin Institute of Technology, in 1998. From 1998 to 2001, she was a Lecturer with the Department of Electrical Engineering, North China Electric Power University, and from 2001 to 2007, she was an Associate Professor. Since 2007, she has been a Professor with the School of Electrical and Electronic Engineering. Her current research interests are in the areas of



**HUITING ZHANG** (Student Member, IEEE) received the B.S. degree from the Taiyuan University of Technology, Taiyuan, China, in 2017. She is currently pursuing the M.S. degree in power systems and automation with North China Electric Power University, Beijing, China. Her research interests include grid connected technology of new energy generation and power system fault diagnosis based on deep reinforcement learning.