

Received August 14, 2020, accepted September 4, 2020, date of publication September 9, 2020, date of current version October 1, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3022890

# Distributed Operation of Wind Farm for Maximizing Output Power: A Multi-Agent Deep Reinforcement Learning Approach

VAN-HAI BUI <sup>ID</sup>, (Member, IEEE), THAI-THANH NGUYEN <sup>ID</sup>, (Member, IEEE),  
AND HAK-MAN KIM <sup>ID</sup>, (Senior Member, IEEE)

Department of Electrical Engineering, Incheon National University, Incheon 406-840, South Korea  
Research Institute for Northeast Asian Super Grid, Incheon National University, Incheon 22012, South Korea

Corresponding author: Hak-Man Kim (hmkim@inu.ac.kr)

This work was supported by Korea Electric Power Corporation under Grant R18XA03.

**ABSTRACT** In the conventional operation of a wind farm (WF) system, the operation point of each wind turbine generator (WTG) is determined to capture maximum energy individually using maximum power point tracking (MPPT) algorithm. However, this operation strategy might not ensure the maximum output power of WF due to wake effect among WTGs. Therefore, this paper develops a multi-agent-based cooperative learning strategy among WTGs using deep reinforcement learning to enhance the overall efficiency of WF by minimizing the wake effect. WTG agents are learnable units and they interact with others as an extensive-form game based on a cooperative model to achieve a common goals (i.e. maximum output power of the WF). In this game, WTG agents carry out their actions sequentially and measure a common reward which is used to update the knowledge of all agents. During the training process, WTG agents use different deep neural networks (DNNs) to improve their actions for achieving the higher reward in the long run by optimizing the weights of DNNs in each learning step. After the training process, WTG agents are able to determine optimal set-points with different input information to minimize the wake effect and to maximize the output power of the WF. Moreover, an operation strategy for the entire WF system is proposed to ensure that the WF always complies with grid-code constraints from transmission system operators, including the requirement of limited power and reserve power. In order to show the effectiveness of the proposed method, a comparison between the results using the proposed method and the conventional MPPT method is also presented in different cases, and the results show that the proposed method can increase the output power of the WF in the range of 1.99% to 4.11% with different layouts.

**INDEX TERMS** Cooperative model, decentralized operation, double deep Q-learning, grid-code constraints, multi-agent deep reinforcement learning, optimization, wind farm operation.

## NOMENCLATURE

### ABBREVIATIONS

DNN	Deep neural network
EMS	Energy management system
MADRL	Multi-agent deep reinforcement learning
MPPT	Maximum power point tracking
TSO	Transmission system operator
WTG	Wind turbine generator
WF	Wind farm

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

### IDENTIFIERS AND SETS

t	Time interval identifier
n	WTG identifier
T	Set of time intervals
N	Set of WTGs

### CONSTANTS AND VARIABLES

$P_{n,t}$	Output power of WTGn at t
$P_{n,t}^{MPPT}$	Output power of WTGn with MPPT at t
$P_n^{rate}$	Rated power of WTGn
$v_{n,t}$	Wind speed at WTGn at t
$\beta_{n,t}$	Pitch angle of WTGn at t

$\lambda_{n,t}$	Tip speed ratio of WTGn at t
$C_P(\beta, \lambda)$	Power coefficient function
$C_T(\beta, \lambda)$	Thrust coefficient function
$P_{WF,t}$	Total output power of WF system
$D_t$	Power requirement from TSO at t
$P_t^{limit}$	Limited power from TSO at t
$p_t^{reser}$	Required reserve capacity from TSO at t
$(s_k, a_k, r_k, s_{k+1})$	Agent transition with information of state, action, reward, and next state at a time step k
$Q(s, a)$	Q-value of a state-action pair $(s, a)$
$L(\theta)$	Loss function with current weights $\theta$
$\theta_i, \theta_{i-1}$	Weights of Q-network and target-network

## I. INTRODUCTION

Due to environmental concerns and exhausting fossil fuels, renewable energy sources such as wind energy, solar energy, hydro energy, etc. have emerged as a new paradigm to fulfill the global energy demand. Among the renewable energy sources, wind energy has attracted significant attention due to its abundant resources [1]. According to the Global Wind Energy Council, the total global installed capacity of wind power was up to 486.8GW by the end of 2016 [2], and a report in [3] predicts that wind power could account for 19% of the worldwide production of electricity by 2030.

The high penetration level of wind energy in the power system has resulted in the development of wind farm (WF) systems with a huge number of wind turbine generators (WTGs). Typically, WTGs are operated at maximum power point tracking (MPPT) to optimize their output power [4]. This operational strategy only maximizes the output power of several WFs where the WTGs are scattered in a large area. Due to the far distance among WTGs, the operation point of WTGs do not affect each other. However, many WF systems are located in an area with limited space among WTGs. Therefore, the operation point of the upstream WTGs can significantly affect the output power of the downstream WTGs due to the decrease in wind velocity [5], [6]. The wind velocity deficit is caused by wake effect in WF, which can reduce the efficiency of the entire WF system by 10-20% [7].

The existing solutions for reducing wake effects in WF system can be categorized into two types, such as optimal placement of WTGs in WF [8], [9], and coordinated operation of WTGs to reduce aerodynamic losses [10], [11]. The first solution is realized by optimal design of WF system, while the second solution is determined by the WF operators during the operation of WF system. It is obvious that coordinated operation of WTGs is necessary after optimizing the layout of the WF system. The main idea of this coordinated operation is that the upstream WTGs might slightly reduce the output power to decrease the wind speed deficit at the downstream WTGs. Therefore, the downstream WTGs can increase their output power, and the total output power of WF might be higher than in the case of operating at conventional MPPT

for all WTGs. In order to achieve this goal, different optimal operation strategies have been proposed to determine the optimal set-point for each WTG by tuning its pitch angle and tip speed ratio [12]–[17].

The authors in [12] have developed a control algorithm for WTGs to increase the overall efficiency of a row of WTGs. In [12], a recursive model has been proposed, which is dependent on thrust coefficients of WTGs, and a centralized control of the WF determines the same set-point for all WTGs. Similar to the conventional MPPT, the same set-points for all WTGs might not ensure the maximum output power of the WF system. The authors in [13] have presented a technical report for the TOPFARM project. This report showed optimal topology design and control of WF system to enhance the overall efficiency of the WF system by adjusting the pitch angle as well as the tip speed ratio. However, the detailed model and result analysis were not presented [14].

The detailed wake model and optimization model have been presented using different optimization methods, such as genetic algorithm [14]–[16] and Adam optimization [17]. The authors in [14] and [15] have developed an optimization model using genetic algorithm to determine the optimal operation point (i.e. tip speed ratio and pitch angle) of each WTG for maximizing the overall WF production. The authors in [17] have proposed a wake steering control algorithm. This algorithm aims to maximize the output power of a WF through yaw misalignment that deflects wakes away from downstream WTGs.

However, these studies [14]–[17] have focused on developing a centralized management system. This means that a centralized energy management system (EMS) gathers all information in the WF system, carries out optimization, and sends optimal set-point to each WTG. Therefore, this EMS requires a two-way communication system [18]. In many WF systems with a large number of WTGs, the communication network for such WF systems will be very complex and it will significantly increase computational burden on the centralized EMS [19]. Moreover, according to [20], [21], utilizing a centralized EMS in a large area, such as WF systems, results in difficulties in management and maintenance of the system. Additionally, these studies [14]–[17] have not considered grid-code constraints from transmission system operators (TSOs). This can lead to several negative impacts on the security and stability of the power system [22], [23]. Therefore, an operation strategy for the whole WF system is required to ensure that the WF system always complies with all the different requirements from TSO.

In order to address the aforementioned issues, this study develops a multi-agent deep reinforcement learning (MADRL)-based operation strategy to enhance the overall efficiency of the WF system. Instead of using a centralized EMS, we develop a multi-agent system that allows the close WTGs (i.e. WTGs in a cluster) to interact with each other and select actions by themselves. This helps to reduce the complexity in the communication network as well as to reduce the computation burden in the WF system.

Furthermore, the cooperative model-based operation of WTGs is also proposed to increase the total output power of the WF system by minimizing the wake effect. The WTG agents are learnable units and they interact with each other in the same environment as an extensive-form game to learn how to select an optimal action in the long run. Each WTG agent uses deep reinforcement learning (DRL) to update the environment information. Therefore, after the learning process, the WTG agents are able to determine their optimal set-points to maximize the output power of the WF. Additionally, an operational strategy for the entire WF system is also proposed, which helps the WF system to comply with various grid-code constraints from TSO such as the constraints of limited power and reserve power. Finally, in order to show the effectiveness of the proposed method, a comparison between the proposed method and the conventional MPPT approach is also presented for different cases. The major contributions of this study are listed as follows.

- A MADRL-based operation strategy is developed to enhance the overall efficiency of a WF by reducing wake effects. Using deep neural networks help to determine the optimal set-points with uncertainty of input information without re-optimization.
- A decentralized management system is developed to reduce both the complexity of the communication network and the computation burden on the system.
- An operational strategy for the entire WF system is also proposed, which helps the WF system to comply with various grid-code constraints from TSO, including the limited power and reserve power constraints.

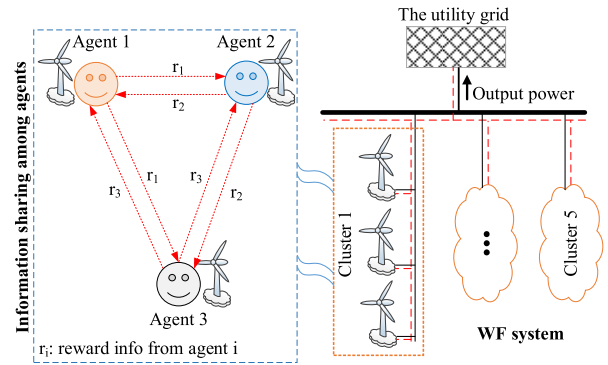
**II. SYSTEM MODEL**

**A. CONFIGURATION OF WIND FARM SYSTEM**

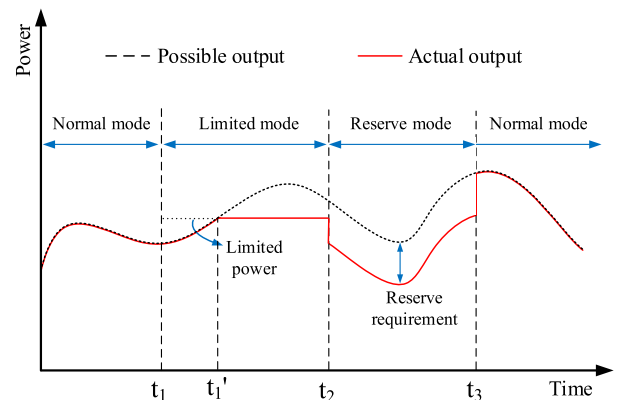
In this study, a WF system including 15 WTGs is used to evaluate the proposed operation strategy. The configuration of the WF system is shown in Figure 1. WTGs are arranged in 5 clusters and the set-point of the WTGs are determined in a decentralized manner. WTGs in a cluster can share information about their rewards with each other. After a training process, each WTG agent is able to select the optimal decisions by itself to reduce the wake effect in the WF system and thereby maximizing the output power of the WF system. In normal operation mode, the WF system always generates maximum power and injects it into the power system. However, TSO may impose several constraints for the WF system in some specific conditions, which are so-called grid-code constraints. During these conditions, the required output power of the WF system and the set-point of WTGs will be changed to satisfy these constraints from TSO. In the next section, different grid-code constraints from TSO will be presented in detail.

**B. INTRODUCTION TO GRID-CODE CONSTRAINTS**

In this section, several grid-constraints are introduced in the operation of the WF system. These constraints are imposed by



**FIGURE 1. Configuration of wind farm system and communication among agents in a cluster.**



**FIGURE 2. Grid-code constraints for the operation of wind farm.**

TSO to ensure that the operation of the WF system does not affect the stability of the power system as well as to support the power system in emergency cases [22]–[24]. There are two grid-code constraints that are often considered in the operation of WF systems, including the requirements for limited power and reserve power [23], [24]. Firstly, the output power of the WF system is bounded by a maximum output power (i.e. the limited power). This means if the total output power of WF is less than the limited power, the WF system is set to generate maximum output power to the power system, as shown in Figure 2 from  $t_1$  to  $t_1'$ . Conversely, if the total output power of WF is greater than the limited power, the set-point of WTGs in the WF system need to reschedule to ensure that the output power of WF equal to the limit power, as shown in Figure 2 from  $t_1'$  to  $t_2$ . Secondly, if the WF system operates in reserve power mode, the set-point of the WTGs also need to reschedule to maintain a certain reserve capacity in the WF system, as shown in Figure 2 from  $t_2$  to  $t_3$ . By using the proposed operation strategy, the WF system can adjust its output power to satisfy various requirements from TSO.

**C. COOPERATIVE MADRL MODEL-BASED OPERATION OF WTGs**

**1) INTRODUCTION TO MULTI-AGENT DEEP REINFORCEMENT LEARNING**

MADRL is a system of agents interacting with each other in a common environment. Each agent performs an action

in every time step along with other agents to complete a given task. Generally, these agents are learnable units that use a particular learning method (i.e. RL or deep RL). Each agent can operate independently to maximize its reward (i.e. competitive model), or agents can also work together to learn a policy for maximizing a common reward in the long run (i.e. cooperative model) by interacting within the same environment [25]–[27].

In a WF system, the major task for the WTGs is to maximize the output power of the entire WF system. Therefore, the reward of a single agent is not much important, WTG agents need to work cooperatively to generate the maximum output power of the WF system. Initially, the upstream WTG agent performs an action to determine its operation point, then the next WTG agents will alternately perform their actions. After that, a common reward will be calculated and this information is used to update the knowledge of all agents to maximize the common reward in the long run. This process is an extensive-form game among agents with a common goal [27], as shown in Figure 3.

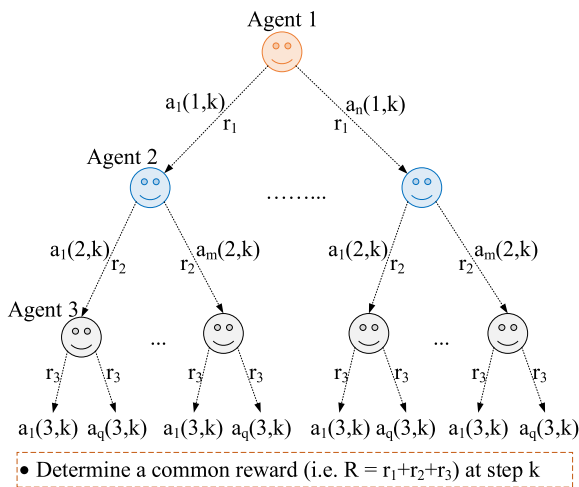


FIGURE 3. Cooperative model-based operation of WTG agents.

Therefore, this study develops a cooperative model-based operation of WTG to determine the set-points of WTGs using multi-agent deep reinforcement learning in a decentralized way. This allows the WTG agents to cooperate in the same environment to maximize the total output power of the WF system. A reward function is designed in the WF system and a detailed operation strategy for all WTG agents also presented in details in the following section.

2) REWARD FUNCTION AND COOPERATIVE MADRL-BASED OPERATION STRATEGY

a: REWARD FUNCTION

Algorithm 1 presents a reward function for the WTG agents in a time step. Initial states  $s$  of WTG agents are taken as input data. Each agent then alternately selects and carries out their actions. After performing this process, each agent receives a reward and observes a new state  $s'$ . These agents'

Algorithm 1 Reward Function in a Learning Step

```

input:  $s = [v, \beta, \lambda]$ 
for agent = 1:M do
    measure output power ( $r$ ) in current state
    select action  $a$  using  $\epsilon$ -greedy policy
    carry out action  $a$ 
    update value of  $(\beta, \lambda)$ 
    measure output power ( $r'$ ) with  $(\beta', \lambda')$ 
    observe next state  $s' = [v', \beta', \lambda']$  and  $R_m = r' - r$ 
    update state  $s = s'$ 
end
    
```

Calculate a common reward for all agents:  $R = \sum_{m=1}^M R_m$

reward information is used to calculate the common reward and update the knowledge of agents (i.e. Q-tables or DNNs).

b: COOPERATIVE MADRL-BASED MODEL

According to the discussion in the previous section, in order to maximize the capacity of the entire WF system, WTG agents need to cooperate in an extensive-form game in a decentralized way. In this study, double deep Q-learning is used for each WTG agent, which helps the agents are able to learn from its experience and improve their actions with experience replay. The detailed learning process is presented in Algorithm 2. Firstly, replay memory size, mini-batch size, and the weights of DNNs for each agent are initialized. Each WTG agent determines the current state information including wind speed ( $v$ ), pitch angle ( $\beta$ ), and tip speed ratio ( $\lambda$ ), then performs an action and observes a reward and new state. All transitions ( $s_k, a_k, r_k, s_{k+1}$ ) are stored in the replay memory for experience replay. In each training step, a mini-batch is randomly drawn from the replay memory and is used to train the Q-network to minimize the mean squared error. The weights of target-network are replaced with the parameters of Q-network after each constant time steps  $C$ . After completing the learning process with a large number of episodes, WTG agents are able to select optimal actions using its DNN and maximize the common reward of the whole system (i.e. maximum total output of WF system).

D. DETAILED OPERATION STRATEGY FOR WF

Figure 4 shows the proposed operation strategy for WF in detail. Initially, input data such as wind speed, initial value of pitch angle, tip speed ratio, and so on are gathered at each WTG. This information is used for WTG agents during the training process. In each training step, an upstream agent (i.e. WTG1) determines an initial state based on the input data and selects an action using epsilon-greedy policy and DNNs. The agent then carries out the selected action, obtain a reward, and a new state. Similarly, the same process is performed by all agents (WTG2 to WTGn) and then a common reward is calculated, as shown in Algorithm 1. This common reward is used to update the weights of DNNs of all WTGs in each training step. The WTG agents cooperate to maximize

**Algorithm 2** Cooperative MADRL-Based Operation Strategy

```

initialize replay memory D and mini-batch size
initialize Q-network and target-network with random
weights  $(\theta, \theta')$ 
for episode = 1, E do:
  for k = 1, K do:
    initialize initial state  $s_m = [v_m, \beta_m, \lambda_m]$ 
    for m = 1, M do:
      select an action at  $s_{m,k}$  using  $\epsilon$ -greedy policy
      carry out  $a_{m,k}$ , achieve  $r_{m,k}$  and next state
       $s_{m,k+1}$ 
      update state  $s_m = s_{m,k+1}$ 
    end
    store transition  $(s_{m,k}, a_{m,k}, \sum_{m=1}^M r_{m,k}, s_{m,k+1})$  in D
    sample random mini-batch of transitions
     $(s_j, a_j, r_j, s_{j+1})$  from D
    estimate target  $y_j$ :
    
$$y_j = \begin{cases} r_j & \text{if terminal state} \\ r_j + \gamma Q(s_{j+1}, \arg \max_a (Q(s_{j+1}, a; \theta)); \theta') & \text{otherwise} \end{cases}$$

    update  $\theta$  using gradient descent to minimize loss
     $(y_j - Q(s_j, a_j; \theta))^2$ 
    reset  $\theta' = \theta$  after C steps
  end
end
  
```

the common reward for the operation of WF during a large number of episodes, as shown in Algorithm 2 in detail.

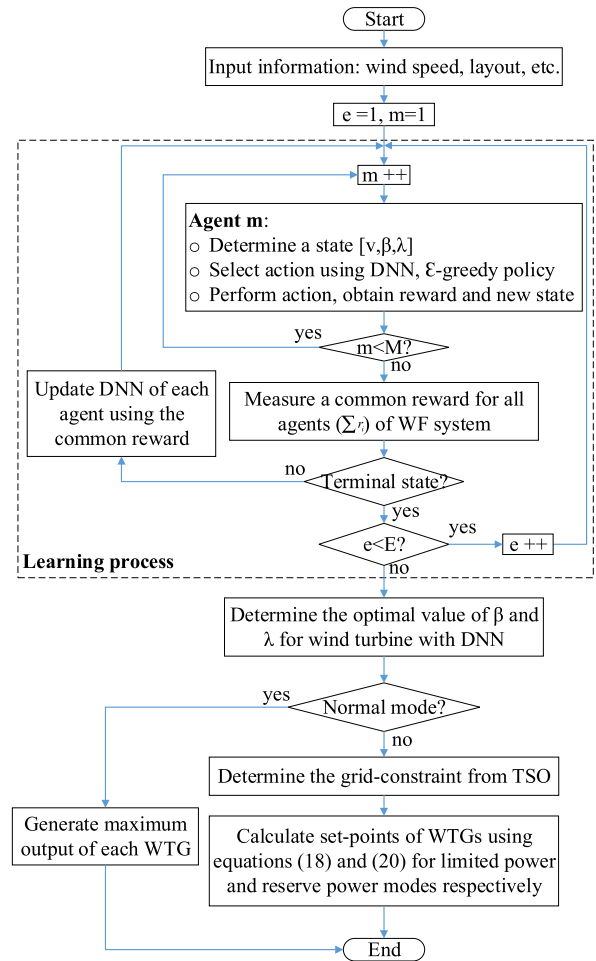
After completing the training process, each WTG agent can determine its optimal set-point using DNNs with optimal parameters. During the operation time, WTG agents also update the information about its operation mode from TSO (i.e. grid-code constraints). If the WF system is operated in normal mode, all WTGs generate maximum output power, which is determined by the training process. By contrast, if there is any grid-code constraint from TSO, the set-point of WTGs are determined using equations (18) and (20) for limited power and reserve power modes, respectively.

**E. MATHEMATICAL MODEL**

Firstly, a brief background of RL including Q-learning and deep Q learning is presented in this section. Secondly, the mathematical model for WTGs and the WF system is also presented in detail considering wake effect in the operation of the WF system. Finally, the set-point of WTGs are determined to fulfill the power requirement from TSO in the different operation modes of the WF system.

**1) BACKGROUND OF REINFORCEMENT LEARNING**

Recently, reinforcement learning (RL) has been widely applied in power system and smart grids [28], [29]. A RL agent is modeled to carry out sequential decision-making by interacting with a particular environment. The environment



**FIGURE 4.** Proposed operation strategy for the wind farm system.

is typically stated in the form of a Markov decision process (MDP), which is defined by a tuple  $(S, A, P_a(s, s'), R_a(s, s'), \gamma)$ , where  $S$  is a set of agent states,  $A$  is a set of actions of the agent,  $P_a(s, s')$  is transition probability from  $s \in S$  to  $s' \in S$  under action  $a \in A$ ,  $R$  is immediate reward by transition  $s$  to  $s'$  with action  $a$ , and  $\gamma \in [0, 1]$  is discount factor for trade-off between immediate rewards and future rewards. The purpose of solving the MDP is to find a policy  $\pi$  that maps the state space  $S$  to a distribution over the action space  $A$  to maximize the discounted accumulated reward.

A popular value-based RL method, so-called Q-learning has been widely applied for optimal operation and control of smart grids or microgrids. This method is to determine an estimate of the Q-value function  $Q(s, a)$ . Whenever the agent carries out a transition  $(s, a, s', r)$ , the Q-value for the state-action pair  $(s, a)$  is updated, as expressed in equation (1).

$$Q(s, a) \leftarrow Q(s, a) + \alpha \cdot \left[ r + \gamma \cdot \max_{a'} Q(s', a') - Q(s, a) \right]$$

$$s \leftarrow s' \tag{1}$$

In Q-learning, all Q-values for state-action pairs are stored in a Q-table, which results in a significant increase in size

of Q-table to solve problems with a huge state space or action space [30]–[32]. To address this issue, deep Q-learning (DQL) have been introduced using deep neural networks (DNNs) as function approximators. These networks are used to map input state information to Q-values for all state-action pairs [31]–[33]. In DQL, the weights of DNNs are trained after each step using a mini-batch with size  $J$  that is randomly drawn from a replay memory. The objective of this training process is to minimize loss function  $L(\theta_i)$ , as given in equation (2).

$$L(\theta_i) = \frac{1}{J} \sum_{j=1}^J (y_j - Q(s_j, a_j; \theta_i))^2 \quad (2)$$

$$y_j = \begin{cases} r_j & \text{if } d_j = \text{True} \\ r_j + \gamma \cdot \max_{a'} Q(s', a'; \theta_{i-1}) & \text{otherwise} \end{cases} \quad (3)$$

It can be observed from equations (2) and (3) that target  $y_j$  and  $Q(s, a)$  are estimated separately by target-network ( $\theta_{i-1}$ ) and Q-network ( $\theta_i$ ). The weights of the target-network is replaced by the weights of Q-network every C iterations.

## 2) WIND TURBINE AND WAKE EFFECT

In this section, the model for WTG and wake effect in the operation of WF will be presented in detail. The output power of each WTG is determined by equation (4).

$$P_{n,t} = \begin{cases} 0 & v_{n,t} < v_{cut-in} \text{ OR } v_{n,t} > v_{cut-out} \\ \frac{1}{2} C_p(\beta, \lambda) \rho \pi R^2 v_{n,t}^3 & v_{cut-in} \leq v_{n,t} < v_{rate} \\ P_n^{rate} & v_{rate} \leq v_{n,t} \leq v_{cut-out} \end{cases} \quad \forall n \in N, t \in T \quad (4)$$

It can be seen from equation (4) that the output power of each WTG  $P_{n,t}$  mainly depends on wind speed  $v_{n,t}$ , and the power coefficient  $C_p$ . In order to increase its output power, each WTG needs to adjust its operation point to maximize the value of  $C_p$ . The value of  $C_p$  with different values of pitch angle ( $\beta$ ), and tip speed ratio ( $\lambda$ ) is shown in Figure 5 for the NREL 5 MW reference WTG [34]. It can be observed that the maximum value of  $C_p = 0.487$  at  $\beta = 0$ , and  $\lambda = 7.6$ . WTGs are generally operated at MPPT with the optimal value of  $C_p$  and the total output power of the WF system is determined by equation (5).

$$P_{WF,t}^{max} = \sum_{n=1}^N P_{n,t}^{MPPT} \quad \forall n \in N, t \in T \quad (5)$$

This operation strategy helps to capture maximum power in WF systems with a far distance among WTGs and therefore, the operation of WTGs do not affect each other. However, due to limited availability of land, WTGs are placed closer to reduce the size of the WF system. Therefore, the operation of downstream WTGs are highly affected by the operation of upstream WTGs due to wake effect.

The wake effect results in a decrease in wind speed at the downstream WTGs. To analyze the effects of wake effect, two

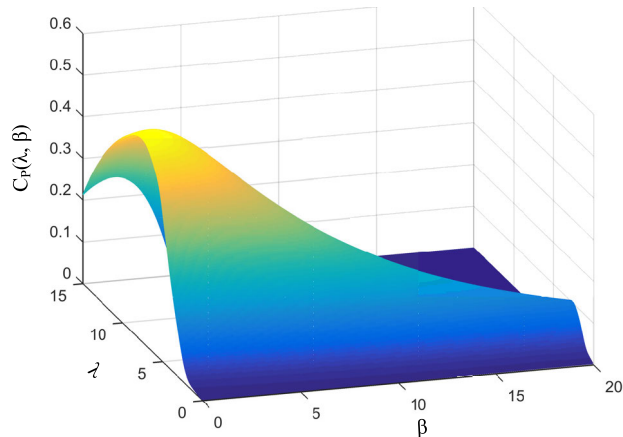


FIGURE 5. Power coefficient ( $C_p$ ) of the NREL 5 MW WTG with different values of pitch angle ( $\beta$ ), and tip speed ratio ( $\lambda$ ).

WTGs are assumed to be placed at a distance  $d$  and are operated at MPPT. This means that WTGs are adjusted to maximize their output power, as shown in equations (6) and (7).

$$P_1 = \frac{1}{2} C_p^{max} \rho \pi R^2 v_0^3 \quad (6)$$

$$C_p^{max} = C_p(\beta = 0, \lambda = 0.76) = 0.487 \quad (7)$$

$$v_2 = v_0 \cdot \left( \frac{1}{2} + \frac{1}{2} \sqrt{1 - 2C_T(\beta_1, \lambda_1) \cdot \left( \frac{D_0}{D_d^w} \right)} \right) \quad (8)$$

$$D_d^w = D_0 + 2k \cdot d \quad (9)$$

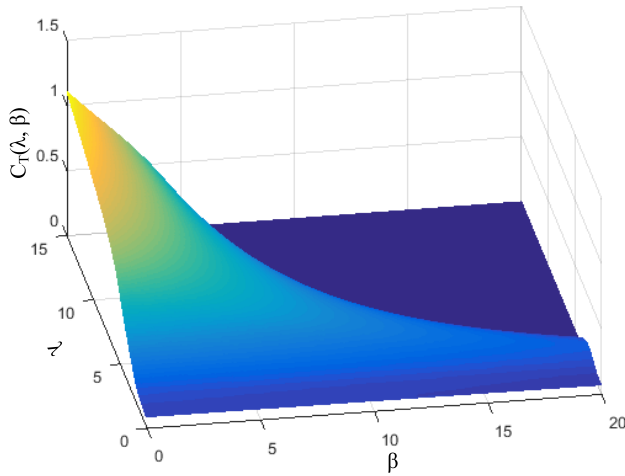
$$k = \begin{cases} 0.075 & \text{onshore wind farm} \\ 0.05 & \text{offshore wind farm} \end{cases} \quad (10)$$

The wind speed at WTG2 is calculated using equations (8)–(10) [14], [34] considering wake effect, where  $D_0$  is the diameter of the rotor,  $d$  is the distance between two WTGs,  $k$  is the entrainment constant,  $k = 0.075$  for onshore WF and  $k = 0.05$  for offshore WF,  $C_T(\beta, \lambda)$  is thrust coefficient. In the operation of WF, the value of  $C_T$  directly affects the velocity at the downstream WTGs, as shown in equation (8). The relationship between  $C_T$ ,  $\beta$ , and  $\lambda$  is shown in detail in Figure 6 [34]. It can be observed that the decrease in wind speed at the downstream WTG is proportional to the value of  $C_T$ . Therefore, in order to reduce the wind speed deficit, it is necessary to reduce the value of  $C_T$  by adjusting the value of  $\beta$  and  $\lambda$ .

$$P_2 = \frac{1}{2} C_p^{max} \rho \pi R^2 v_2^3 \quad (11)$$

$$P_{output}^{WTGs} = \frac{1}{2} C_p^{max} \rho \pi R^2 (v_0^3 + v_2^3) \quad (12)$$

The output power of the downstream WTG and total output power of WTGs are presented in equations (11) and (12) respectively. It can be seen that the operation of each WTGs at MPPT cannot guarantee the maximum output power of the WF system. This is because there does not exist a pair of ( $\beta$ ,  $\lambda$ ) to maximize the value of  $C_p$  and minimize the value of  $C_T$ . Therefore, it is possible to find an optimal pair of ( $\beta$ ,  $\lambda$ ) for



**FIGURE 6.** Thrust coefficient ( $C_T$ ) of the NREL 5 MW WTG with different values of pitch angle ( $\beta$ ), and tip speed ratio ( $\lambda$ ).

each WTG to maximize the output power of the whole WF system, and these set-points may differ from MPPT. In this paper, a cooperative model will be developed to find these values using MADRL-based operation strategy. Objective function and different operation constraints are expressed in equations (13)-(15).

$$\text{Max} \left\{ P_t^{WF} = \sum_{n=1}^N P_{n,t}^{WTG} \right\} \quad (13)$$

$$\beta_n^{\min} \leq \beta_{n,t} \leq \beta_n^{\max} \quad \forall n \in N, t \in T \quad (14)$$

$$\lambda_n^{\min} \leq \lambda_{n,t} \leq \lambda_n^{\max} \quad \forall n \in N, t \in T \quad (15)$$

### 3) SET-POINT OF WTGs WITH GRID-CODE CONSTRAINTS

In normal operation mode, WF is required to generate the maximum output power and injects into the power system. However, the required power from TSO might change depending on the operation conditions of the power system, as discussed in section II. B. There are two common grid-code constraints for the operation of WF system, including limited power and reserve power modes [22]–[24]. In limited power mode, the required power is determined by equation (16). It can be seen that the amount of output power of a WF system will always be bounded by a constant value of limited power. Besides, if WF operates in reserve power mode, the amount of output power is determined based on the required reserve capacity, as given in equation (17).

$$D_t = \begin{cases} \sum_{n=1}^N P_{n,t}^{\max} & \text{if } \sum_{n=1}^N P_{n,t}^{\max} < P_t^{\text{limit}} \\ P_t^{\text{limit}} & \text{otherwise} \end{cases} \quad \forall t \in T \quad (16)$$

$$D_t = (1 - p_t^{\text{reser}}) \cdot \sum_{n=1}^N P_{n,t}^{\max} \quad \forall t \in T \quad (17)$$

After determining the required output power for the WF system, the set-point of WTGs are also determined to fulfill

the required power from TSO in different operation modes. In limited power mode, if the amount of power limit is greater than the maximum output power of the WF system, the set-point of WTGs is set to generate the maximum power. By contrast, if the limit power output is less than the maximum output power of WF, the set-point of WTGs need to reduce to balance the required power from TSO and the amount of power reduction for each WTG is calculated based on the amount of mismatch power and the number of WTGs, as shown in equations (18) and (19). In reserve power mode, the set-point of WTGs are determined simply by maintaining the same proportion of reserve capacity for each WTGs, as given in equation (20).

$$P_{n,t} = \begin{cases} P_{n,t}^{\max} & \text{if } \sum_{n=1}^N P_{n,t}^{\max} \leq P_t^{\text{limit}} \\ P_{n,t}^{\max} - \frac{DP_t}{N_{WTG_s}} & \text{otherwise} \end{cases} \quad \forall n \in N, t \in T \quad (18)$$

$$DP_t = \left( \sum_{n=1}^N P_{n,t}^{\max} - P_t^{\text{limit}} \right) \quad \forall t \in T \quad (19)$$

$$P_{n,t} = P_{n,t}^{\max} \cdot (1 - p_t^{\text{reser}}) \quad \forall n \in N, t \in T \quad (20)$$

## III. NUMERICAL RESULTS

In this section, the training process is presented in detail to determine the optimal set-point of WTGs in a decentral-ized manner. Moreover, several grid-code constraints also are applied to the operation of the WF system.

### A. INPUT DATA

In this study, a test WF system consisting of 15 WTGs divided into 5 clusters is used to evaluate the proposed method, as shown in Figure 1. The detailed parameters for a WTG are as follows [34].

- Rated power: 5 MW
- Rotor diameter ( $D_0$ ): 63m
- Rho ( $\rho$ ): 1.2231kg/m<sup>3</sup>
- Distance between two consecutive WTGs ( $d$ ):  $5D_0 = 315\text{m}$
- $v_{\text{cut-in}} = 3\text{m/s}$  and  $v_{\text{cut-out}} = 25\text{m/s}$

### B. TRAINING PROCESS AND OPTIMAL RESULTS

The MADRL-based model was trained with 10000 episodes. The learning rate  $\alpha = 0.1$ , discount factor  $\gamma = 0.999$ . The initial value of pitch angle  $\beta$  and tip speed ratio  $\lambda$  is 10.0 and 5.0, respectively. The wind speed at the upstream WTG is assumed at 12m/s.

The value of epsilon during the training process is shown in Figure 7. This value was reduced by an epsilon decay to ensure the trade-off between exploitation and exploration for each agent during the training process. At the end of the training process, the value of epsilon decreased to the minimum value and agents mainly select an action based on its knowledge about the environment (i.e. DNNs with optimal

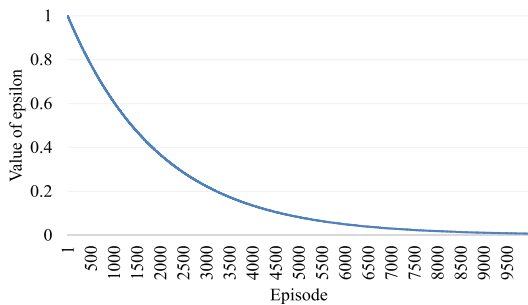


FIGURE 7. Value of epsilon during the training process.

parameters). In order to guarantee the acceptable accuracy, WTGs might be trained during few hours. However, this training process is off-line training based on different input data. After training, WTGs can use the optimal DNNs to estimate the Q-values of possible actions, and then determine the optimal set-point without re-optimization.

Additionally, the total reward for each agent is shown in Figure 8. It can be noticed that the total reward ( $\sum r_i = r_1 + r_2 + r_3$ ) converges to the optimal value. This is because agents cooperate to maximize the total reward in the long run. At the beginning of the training process, agents often choose randomly actions to explore the environment by setting a high value of epsilon. After having enough information about the environment, agents mainly choose actions using their DNN to increase the total reward during the training process.

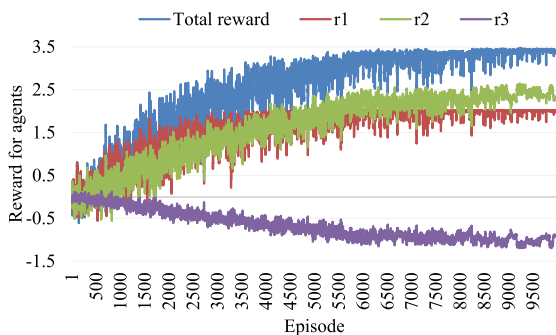


FIGURE 8. The total reward and single reward of each WTG agent.

After the training process, agents can use DNNs to select optimal actions to increase the total output power of the WF system. In this section, an example is presented to show how the WTG agents in a cluster select their actions in a given state using DNN with optimal parameters. The wind speed at the upstream WTG is 12m/s in this test case. The initial state of agent1 (WTG1) is  $[v_0, \beta, \lambda] = [12.0, 10.0, 5.0]$ . The Q-values for all possible state-action pairs are shown in Figure 9. The best action in this state is the action assigned to the highest Q value, i.e. decrease the value of  $\beta$  and increase the value of  $\lambda$ . After performing a series of actions, it is easy to find the optimal operation point of WTG1 at  $\beta_{opt} = 2$  and  $\lambda_{opt} = 5.2$ .

Similar to the agent1, Q-values for all possible state-action pairs of agent2 are shown in Figure 10. At the initial state  $[v_0,$

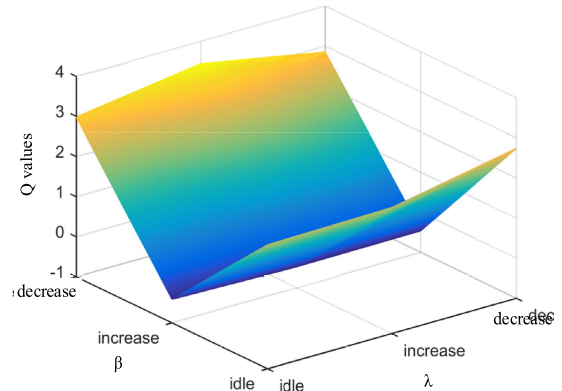


FIGURE 9. Q-values at the initial state [12.0, 10.0, 5.0] of agent 1.

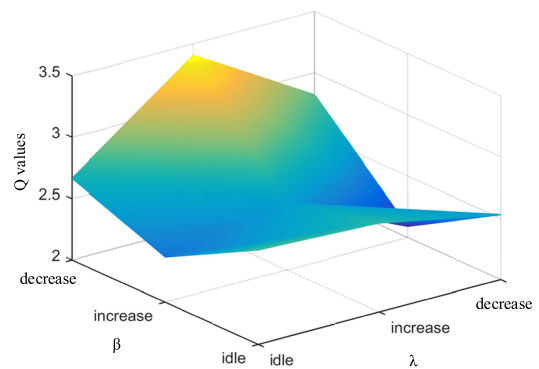


FIGURE 10. Q-values at the initial state [11.4, 10.0, 5.0] of agent 2.

$\beta, \lambda] = [11.4, 10.0, 5.0]$ , the best action is to decrease the value of  $\beta$  and increase the value of  $\lambda$ . Finally, the optimal operation point of WTG2 is determined at  $\beta_{opt} = 0$  and  $\lambda_{opt} = 6.9$ . In this test case, WTG3 is the last WTG in a cluster, so this WTG is set at MPPT ( $\beta_{opt} = 0$  and  $\lambda_{opt} = 7.6$ ) to maximize its output considering the decrease in wind speed by WTG1 and WTG2.

C. COMPARION WITH THE CONVENTIONAL METHOD

In this section, in order to show the effectiveness of the proposed method, the results using the proposed method are compared to the results using the conventional MPPT method with different scenarios.

1) OUTPUT POWER OF A CLUSTER WITH DIFFERENT WIND SPEED

To evaluate the proposed method, wind speed at upstream WTG is varied from  $v_{cut-in}$  to  $v_{cut-out}$ . The distance between two consecutive WTGs is  $5D_0$ . Figure 11 shows the total output power of a cluster with three WTGs by using both the proposed method and conventional method (i.e. MPPT). It is easy to see that if  $v_0 < 3m/s$ , WTGs do not generate power because the wind speed is less than  $v_{cut-in}$  and if  $v \geq 14m/s$ , all WTGs are set to generate the rated output power. If  $v \in [3, 14)$ , there is a difference in the amount of output power using the proposed method and the conventional



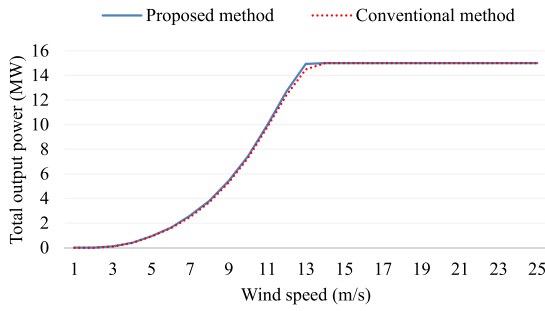


FIGURE 11. Output power of a cluster with different wind speed at upstream WTG.

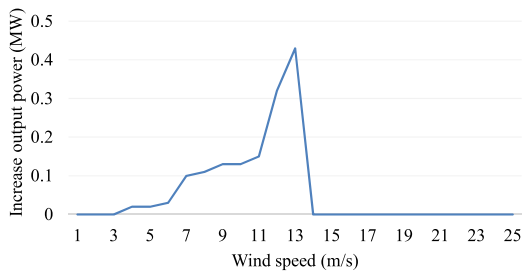


FIGURE 12. Increment of output power using the proposed method.

TABLE 1. Comparison between the proposed method and the conventional MPPT method with different WF layout.

Distance	3D <sub>0</sub>	5D <sub>0</sub>	7D <sub>0</sub>
MPPT mode	10.47	12.4	13.6
Proposed mode	10.9	12.72	13.87
Increase (%)	4.11	2.58	1.99

MPPT method. The difference output power between the two methods is shown in Figure 12. It can be noticed that when the wind speed at upstream WTGs increases from 3m/s to 14m/s, the proposed method can generate more power than the conventional method. In this study, we focused mainly on developing a MADRL-based optimization model for the WF system to improve the overall efficiency of the entire system. Therefore, a simple configuration of a cluster with a row of three WTGs is considered to test the performance of the proposed method, and the output power increases by 2.58%, as shown in Table 1. In the following section, the parameters that affect the amount of increasing power will be analyzed in detail.

## 2) OUTPUT POWER OF A CLUSTER WITH DIFFERENT DISTANCE AMONG WTGS

In this section, the output power of a cluster in the WF system will be analyzed with different distance between WTGs and constant wind speed at upstream WTG, i.e.  $v_0 = 12\text{m/s}$ . Table 1 shows the total output power of a cluster with different

distances between two consecutive WTGs (i.e.  $3D_0$ ,  $5D_0$ ,  $7D_0$ ) using the conventional method (i.e. MPPT) and the proposed method. It can be noticed that the amount of output power increases when the distance between the two WTGs increases. This is because as this distance increases, the wake effect decreases significantly, and therefore, the wind speed at downstream WTGs is not greatly affected by the operation point of the upstream WTGs. Additionally, the increase in the cluster’s output power is 4.11%, 2.58%, 1.99% with different distances  $3D_0$ ,  $5D_0$ , and  $7D_0$ , respectively. This indicates that the wake effect might be negligible if the distance between the WTGs is far. Therefore, the proposed method is used effectively in WFs with short distances between WTGs.

## 3) ANNUAL ENERGY PRODUCTION OF A CLUSTER WITH DIFFERENT METHODS

In order to show the effectiveness of the proposed method more clearly, a comparison of the annual energy production (AEP) between the proposed method and the conventional MPPT method is presented in this section. Wind speed is generally considered to have a certain probability distribution function. In this study, we assume that wind speed follows Weibull Distribution [35], [36], as shown in Figure 13 with detailed parameters given in Table 2, similar to [36]. The wind speed data for a year is generated using the wind speed distribution model in Figure 13, and this information is used to calculate the AEP. The AEP in a cluster is shown in Table 2 using both the proposed method and the conventional MPPT. It can be seen that using the proposed method, the AEP can increase by 989.1MWh.

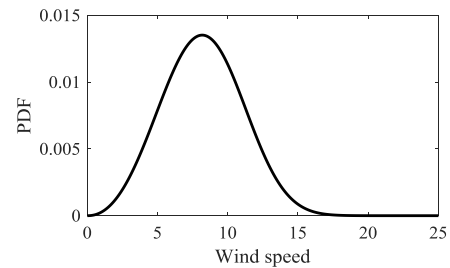


FIGURE 13. Wind speed distribution model.

TABLE 2. Parameters of Weibull distribution model and AEP.

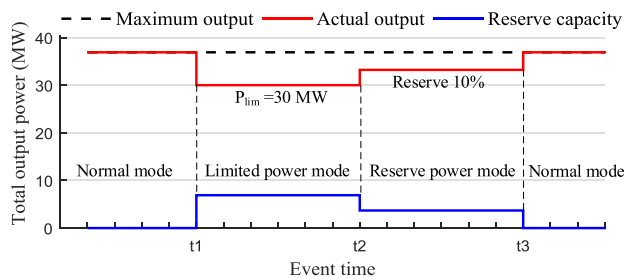
Parameters of Weibull distribution	Mean value of wind speed (m/s)	Weibull scale (m/s)	Weibull form
	8.13	9.2	3.2
AEP	MPPT mode (MWh)	Proposed mode (MWh)	Increasing amount (MWh)
	42834.03	43823.13	989.1

## D. HANDLING VARIOUS GRID-CODE CONSTRAINTS FROM TSO

Generally, WTGs are set to maximize the total output power of the WF system, which might affect the stability of the

**TABLE 3.** Input data of maximum output power in WF.

Parameters	Wind speed at upstream WTG	Pmax1	Pmax2	Pmax3	Total
Cluster 1	8	1.9	1.2	0.7	3.8
Cluster 2	9	2.6	1.7	1.1	5.4
Cluster 3	10	3.7	2.3	1.5	7.5
Cluster 4	12	4.9	4.8	3	12.7
Cluster 5	10	3.7	2.3	1.5	7.5
Output power of WF	-	-	-	-	36.9

**FIGURE 14.** Handling different grid-code constraints from TSO in WF.

power system. Therefore, TSO imposes several grid-code constraints for the operation of WF. In order to evaluate the proposed method for handling the grid-code constraints from TSO, the information about the maximum output power of WTGs in the WF system is assumed and tabulated in Table 3.

As stated in part II. B, there are two grid-code constraints for the operation of WF, named limited power and reserve power modes. In this study, we assume that from  $t_1$  to  $t_2$ , WF is operated in limited power mode with limited power at 30MW, and from  $t_2$  to  $t_3$ , WF is operated in reserve power mode with 10% for reserve capacity. The output power of WF is shown in Figure 14. At  $t_1$ , the output power of WF is reduced by 6.9MW to ensure that the output power is always less than or equal to limited power (30MW). This power reduction acts as reserve power in WF. At  $t_2$ , WF should maintain 10% of the output power for reserve capacity, which is 3.69MW. At  $t_3$ , WF operates in normal mode, therefore, the output power increases to the maximum output power at 36.9MW.

The set-point of WTGs in limited power mode are determined by equations (18) and (19). Detailed information about the set-point of WTGs is presented in Table 4 in limited power mode from  $t_1$  to  $t_2$  and the total output power of WF is 30MW. In reserve power mode, the set-point of WTGs is determined by equation (20). Detail information about the set-point of WTGs are tabulated in Table 5 in reserve power mode from  $t_2$  to  $t_3$  and the output power of WF is 33.21MW.

#### IV. CONCLUSION

In this study, a MADRL-based operation strategy has been developed to enhance the overall efficiency of WF system

**TABLE 4.** Detailed set-point of WTGs in limited power mode.

Parameters	Pset1	Pset2	Pset3	Total
Cluster 1	1.44	0.74	0.24	2.42
Cluster 2	2.14	1.24	0.64	4.02
Cluster 3	3.24	1.84	1.04	6.12
Cluster 4	4.44	4.34	2.54	11.32
Cluster 5	3.24	1.84	1.04	6.12
Output power of WF	-	-	-	30

**TABLE 5.** Detailed set-point of WTGs in reserve power mode.

Parameters	Pset1	Pset2	Pset3	Total
Cluster 1	1.71	1.08	0.63	3.42
Cluster 2	2.34	1.53	0.99	4.86
Cluster 3	3.33	2.07	1.35	6.75
Cluster 4	4.41	4.32	2.7	11.43
Cluster 5	3.33	2.07	1.35	6.75
Output power of WF	-	-	-	33.21

by reducing wake effects. Additionally, a decentralized management system is developed to reduce both the complexity of the communication network and the computation burden on the system. All WTGs in the same cluster interact with each other as an extensive-form game based on a cooperative model to achieve a common goal (i.e. maximum output power of WF). Each WTG agent improves its actions using a DNN and the weights of DNN are updated after every learning step. After the training process, the WTGs agents are able to determine the optimal set-points with different input information to minimize the wake effect, and thereby maximizing the output power of WF. Simulation results have shown that the proposed method can increase the output power of the tested WF system in the range of 1.99% to 4.11%, under different layouts, in comparison to the conventional MPPT approach. Additionally, an operational strategy has been proposed for fulfilling the grid-code constraints from TSO, including the limited power and reserve power constraints.

#### REFERENCES

- [1] V. Yaramasu, B. Wu, P. C. Sen, S. Kouro, and M. Narimani, "High-power wind energy conversion systems: State-of-the-art and emerging technologies," *Proc. IEEE*, vol. 103, no. 5, pp. 740–788, May 2015.
- [2] GWEC, "Global wind report 2016," Global Wind Energy Council, Brussels, Belgium, Tech. Rep., 2017. [Online]. Available: [http://www.gwec.net/wp-content/uploads/2014/10/GWEO2014\\_WEB.pdf](http://www.gwec.net/wp-content/uploads/2014/10/GWEO2014_WEB.pdf)
- [3] GWEC. *Wind Energy Outlook: 2000 Gigawatts by 2030*. Accessed: Oct. 2014. [Online]. Available: <https://gwec.net/wind-energy-outlook-2000-gigawatts-2030/>
- [4] D. Kumar and K. Chatterjee, "A review of conventional and advanced MPPT algorithms for wind energy systems," *Renew. Sustain. Energy Rev.*, vol. 55, pp. 957–970, Mar. 2016.

- [5] J. Tian, D. Zhou, C. Su, F. Blaabjerg, and Z. Chen, "Optimal control to increase energy production of wind farm considering wake effect and lifetime estimation," *Appl. Sci.*, vol. 7, no. 1, p. 65, Jan. 2017.
- [6] E. T. Al-Shammari, S. Shamsirband, D. Petković, E. Zalnezhad, P. L. Yee, R. S. Taher, and Ž. Cojbašić, "Comparative study of clustering methods for wake effect analysis in wind farm," *Energy*, vol. 95, pp. 573–579, Jan. 2016.
- [7] H. Kim, K. Kim, C. Bottasso, F. Campagnolo, and I. Paek, "Wind turbine wake characterization for improvement of the ainslie eddy viscosity wake model," *Energies*, vol. 11, no. 10, p. 2823, Oct. 2018.
- [8] R. Shakoor, M. Y. Hassan, A. Raheem, and Y.-K. Wu, "Wake effect modeling: A review of wind farm layout optimization using Jensen's model," *Renew. Sustain. Energy Rev.*, vol. 58, pp. 1048–1059, May 2016.
- [9] J. Serrano González, M. Burgos Payán, J. M. R. Santos, and F. González-Longatt, "A review and recent developments in the optimal wind-turbine micro-siting problem," *Renew. Sustain. Energy Rev.*, vol. 30, pp. 133–144, Feb. 2014.
- [10] A. Isa, J. Chhor, P. Krajiski, V. Kipke, and C. Sourkounis, "Review on optimal wind farm control techniques and prospects of artificial intelligence," in *Proc. 45th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2019, pp. 2481–2488.
- [11] P. Gebraad, J. J. Thomas, A. Ning, P. Fleming, and K. Dykes, "Maximization of the annual energy production of wind power plants by optimization of layout and yaw-based wake control," *Wind Energy*, vol. 20, no. 1, pp. 97–107, Jan. 2017.
- [12] D. Madjidian and A. Rantzer, "A stationary turbine interaction model for control of wind farms," *IFAC Proc. Volumes*, vol. 44, no. 1, pp. 4921–4926, Jan. 2011.
- [13] G. C. Larsen, H. A. Madsen, N. Troldborg, T. J. Larsen, P. E. Réthoré, P. Fuglsang, S. Ott, J. Mann, T. Buhl, M. Nielsen, and H. Markou. *TOPFARM-Next Generation Design Tool for Optimisation of Wind Farm Topology and Operation*. Accessed: 2011. [Online]. Available: <https://orbit.dtu.dk/en/publications/topfarm-next-generation-design-tool-for-optimisation-of-wind-farm>
- [14] J. S. González, M. B. Payán, J. R. Santos, and Á. G. G. Rodríguez, "Maximizing the overall production of wind farms by setting the individual operating point of wind turbines," *Renew. Energy*, vol. 80, pp. 219–229, Aug. 2015.
- [15] J. S. Gonzalez, M. B. Payan, and J. R. Santos, "Optimum wind turbines operation for minimizing wake effect losses in offshore wind farms," in *Proc. 13th Int. Conf. Environ. Electr. Eng. (EEEIC)*, Nov. 2013, pp. 188–192.
- [16] M. Allagui, M. Abbes, and O. B. K. Hasnaoui, "Optimization of the wind farm energy capture by minimizing the wake effects," in *Proc. 6th Int. Renew. Energy Congr. (IREC)*, Mar. 2015, pp. 1–6.
- [17] M. F. Howland, S. K. Lele, and J. O. Dabiri, "Wind farm power optimization through wake steering," *Proc. Nat. Acad. Sci. USA*, vol. 116, no. 29, pp. 14495–14500, Jul. 2019.
- [18] A. D. Hansen, P. Sørensen, F. Iov, and F. Blaabjerg, "Centralised power control of wind farm with doubly fed induction generators," *Renew. Energy*, vol. 31, no. 7, pp. 935–951, Jun. 2006.
- [19] V.-H. Bui, A. Hussain, W.-G. Lee, and H.-M. Kim, "Hybrid energy management system for operation of wind farm system considering grid-code constraints," *Energies*, vol. 12, no. 24, p. 4672, Dec. 2019.
- [20] Z. Wang, K. Yang, and X. Wang, "Privacy-preserving energy scheduling in microgrid systems," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 1810–1820, Dec. 2013.
- [21] A. Hussain, V.-H. Bui, and H.-M. Kim, "A resilient and privacy-preserving energy management strategy for networked microgrids," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 2127–2139, May 2018.
- [22] M. Tsili and S. Papathanassiou, "A review of grid code technical requirements for wind farms," *IET Renew. Power Generat.*, vol. 3, no. 3, pp. 308–332, Sep. 2009.
- [23] Y.-K. Wu, S.-M. Chang, and P. Mandal, "Grid-connected wind power plants: A survey on the integration requirements in modern grid codes," *IEEE Trans. Ind. Appl.*, vol. 55, no. 6, pp. 5584–5593, Nov. 2019.
- [24] V.-H. Bui, A. Hussain, and H.-M. Kim, "Optimal operation of wind farm for reducing power deviation considering grid-code constraints and events," *IEEE Access*, vol. 7, pp. 139058–139068, 2019.
- [25] A. OroojlooyJadid and D. Hajinezhad, "A review of cooperative multi-agent deep reinforcement learning," 2019, *arXiv:1908.03963*. [Online]. Available: <http://arxiv.org/abs/1908.03963>
- [26] Y. Wang, H. Liu, W. Zheng, Y. Xia, Y. Li, P. Chen, K. Guo, and H. Xie, "Multi-objective workflow scheduling with deep-Q-network-based multi-agent reinforcement learning," *IEEE Access*, vol. 7, pp. 39974–39982, 2019.
- [27] K. Zhang, Z. Yang, and T. Başar, "Multi-agent reinforcement learning: A selective overview of theories and algorithms," 2019, *arXiv:1911.10635*. [Online]. Available: <http://arxiv.org/abs/1911.10635>
- [28] Z. Zhang, D. Zhang, and R. C. Qiu, "Deep reinforcement learning for power system applications: An overview," *CSEE J. Power Energy Syst.*, vol. 6, no. 1, pp. 213–225, Mar. 2020.
- [29] D. Zhang, X. Han, and C. Deng, "Review on the research and practice of deep learning and reinforcement learning in smart grids," *CSEE J. Power Energy Syst.*, vol. 4, no. 3, pp. 362–370, Sep. 2018.
- [30] V.-H. Bui, A. Hussain, and H.-M. Kim, "Q-learning-based operation strategy for community battery energy storage system (CBESS) in microgrid system," *Energies*, vol. 12, no. 9, p. 1789, May 2019.
- [31] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, Oct. 2018.
- [32] Y.-H. Bui, A. Hussain, and H.-M. Kim, "Double deep Q-learning-based distributed operation of battery energy storage system considering uncertainties," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 457–469, Jan. 2020.
- [33] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," in *Proc. 13th AAAI Conf. Artif. Intell.*, Mar. 2016, pp. 2094–2100.
- [34] J. M. Jonkman, S. Butterfield, W. Musial, and G. Scott, "Definition of a 5-MW reference wind turbine for offshore system development," Nat. Renew. Energy Lab., Golden, CO, USA, Tech. Rep. NREL/TP-500-38060, Feb. 2009.
- [35] J. A. Carta, P. Ramírez, and S. Velázquez, "A review of wind speed probability distributions used in wind energy analysis," *Renew. Sustain. Energy Rev.*, vol. 13, no. 5, pp. 933–955, Jun. 2009.
- [36] J. Xu, X. Yi, Y. Sun, T. Lan, and H. Sun, "Stochastic optimal scheduling based on scenario analysis for wind farms," *IEEE Trans. Sustain. Energy*, vol. 8, no. 4, pp. 1548–1559, Oct. 2017.



**VAN-HAI BUI** (Member, IEEE) received the B.S. degree in electrical engineering from the Hanoi University of Science and Technology (HUST), Vietnam, in 2013, and the Ph.D. degree in electrical engineering from Incheon National University (INU), South Korea. He is currently a Postdoctoral Researcher with the Department of Electrical Engineering, INU. His research interests include energy management systems, the applications of reinforcement learning in smart grids, and microgrid operation.



**THAI-THANH NGUYEN** (Member, IEEE) received the B.Sc. degree in electrical engineering from the Hanoi University of Science and Technology, Vietnam, in 2013, and the Ph.D. degree in electrical engineering from Incheon National University (INU), South Korea, in 2019. He is currently a Researcher with Incheon National University (INU), South Korea. His research interests include power system control and stability, power converter control, and the applications of power

electronics to power systems.



**HAK-MAN KIM** (Senior Member, IEEE) received the first Ph.D. degree in electrical engineering from Sungkyunkwan University, South Korea, in 1998, and the second Ph.D. degree in information sciences from Tohoku University, Japan, in 2011. He worked with the Korea Electrotechnology Research Institute (KERI), South Korea, from October 1996 to February 2008. He is currently a Professor with the Department of Electrical Engineering, Incheon National University, South Korea. His research interests include microgrid operation and control, and dc power systems.

• • •