

ORIGINAL RESEARCH

Open Access



Jointly improving energy efficiency and smoothing power oscillations of integrated offshore wind and photovoltaic power: a deep reinforcement learning approach

Xiuxing Yin¹ and Meizhen Lei^{2*}

Abstract

This paper proposes a novel deep reinforcement learning (DRL) control strategy for an integrated offshore wind and photovoltaic (PV) power system for improving power generation efficiency while simultaneously damping oscillations. A variable-speed offshore wind turbine (OWT) with electrical torque control is used in the integrated offshore power system whose dynamic models are detailed. By considering the control system as a partially-observable Markov decision process, an actor-critic architecture model-free DRL algorithm, namely, deep deterministic policy gradient, is adopted and implemented to explore and learn the optimal multi-objective control policy. The potential and effectiveness of the integrated power system are evaluated. The results imply that an OWT can respond quickly to sudden changes of the inflow wind conditions to maximize total power generation. Significant oscillations in the overall power output can also be well suppressed by regulating the generator torque, which further indicates that complementary operation of offshore wind and PV power can be achieved.

Keywords Offshore wind turbine, Offshore photovoltaic power, Deep reinforcement learning, Deep deterministic policy gradient, Multi-objective optimal control

1 Introduction

Decarbonization of electricity production is critical for addressing the issues of climate change and global warming. Improvements in the efficiency and stability of renewable power generation such as wind and photovoltaic (PV) energy enable a more rapid and lower-cost transition to a decarbonized energy system. The joint operation of wind and PV power systems is especially effective in increasing renewable power production [1].

Based on their complementary nature, the integration and coordinated operation of wind and PV power ensure a reliable and adequate power supply, while adhering to the grid requirements by enhancing the peak power shaving capacity. The integration of an offshore wind turbine (OWT) with offshore PV can potentially enhance energy output and reduce overall project cost because of the shared seawater space, power infrastructure and mooring system. In addition, offshore PV can be complementary to OWTs in aspects of increasing the planned cable capacity factor and improving the turbine design life span [2]. Therefore, the integration of offshore wind and PV systems will be beneficial in both economical and technical terms.

Recently, a variety of methods have been investigated, primarily directed at the optimization and coordinated

*Correspondence:

Meizhen Lei
lixingfile@163.com

¹ The State Key Laboratory of Water Resources Engineering and Management, Wuhan University, Wuhan 430072, Hubei, China

² The School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China

operation of complementary wind-PV systems to ensure power generation quality and efficiency. A multi-objective model is established to maximize power generation and minimize output fluctuations of a hydro-wind-photovoltaic power system in [3], whereas in [4], the feasibility of adding an offshore floating solar farm to an existing Dutch offshore wind farm under the constraint of a certain fixed cable capacity is evaluated. The optimal size and operation of complementary hydro-wind-PV power systems are explored to provide reliable and adequate power for the grid in [5], while [6] presents a model for estimating emissions of electricity from systems that couple photovoltaic and wind generation with lithium-ion and vanadium redox flow batteries. In [7], the potential of combining offshore wind and solar power is explored based on the technical specifications of commercial wind turbines and PV panels, while in [8], a two-stage evaluation mode-based fuzzy multi criteria decision making method is proposed to select the optimal site of wind-PV power plants given the different attitudes of decision makers. Reference [9] suggests a multi-objective particle swarm optimization method to enhance the operation of a wind farm and a PV array with a battery energy storage system, while [10] presents the analyses of large-scale integration of wind and PV power into a Danish reference energy system by considering certain ancillary services to identify optimal mixtures from a technical point of view. In [11], it shows that renewables could provide a source of power competitive with fossil-based alternatives in India by using a cost optimization model.

Reference [12] uses a mathematical tool, Copula, to unscramble the dependencies between the power of wind and PV plants and introduces a probability method to analyze how power and energy are compensated at a certain confidence level. In [13], fault detection, classification, and location for a PV-Wind-based microgrid are presented, whereas [14] proposes a damping method for low frequency oscillations by incorporating a supplementary damping controller with a PV generating station whose parameters are coordinated with a power system stabilizer. Reference [15] proposes a 2-degree of freedom combined proportional-integral and derivative control scheme for the frequency and power control of a wind integrated interconnected power system.

However, most recent works mainly focus on the economic/technical planning and feasibility analysis of the onshore wind-PV systems at the planning level, while the real-time operation and optimization of the offshore

integrated wind-PV system to improve energy efficiency and to smooth power oscillations have not been well investigated. There has been limited attention to the multi-objective real-time operations of the integrated offshore wind and PV power system by considering the intermittent and stochastic nature of wind and solar PV facilities, and hence the real-time complementary potential has not been fully explored. In addition, the conventional proportional-integral and derivative control methods, which have been widely employed in the current literature, may not fulfill the high-quality joint control objectives of the optimal wind power capture and power oscillation smoothing for an integrated OWT and PV power system.

This paper proposes a deep reinforcement learning (DRL) approach and deep deterministic policy gradient (DDPG) algorithm for the joint operations of the integrated offshore wind and PV power system. Inspired by behavioral psychology, DRL has exhibited high potential in dealing with sequential complex decision-making problems such that the cumulative reward can be maximized when interacting with an uncertain environment. DRL is also adaptive and model-free, and does not need prior knowledge of the environment, as it can learn the generalized optimal control strategy from historical data. By using DDPG for real-time control, the OWT rotor speed can be varied according to inflow wind conditions while the generator rotation speed can be synchronized with the grid frequency, and the complementary operation and oscillation suppression of the OWT and the photovoltaic power can be well achieved. The potential and effectiveness of the integrated OWT and PV power system are evaluated based on design experiments.

The main novelties and contributions of this work are as follows:

- (a) A control framework for the real-time operation and optimization of an offshore integrated wind-PV system to jointly improve energy efficiency and smooth power oscillations.
- (b) The DDPG methodology design for achieving a tradeoff between the power generation and power oscillation damping performance of the integrated OWT-PV power system.
- (c) Verifications of the DDPG approach in simultaneously improving the power capture efficiency and power smoothing capability of the integrated power system using only the generator torque control.

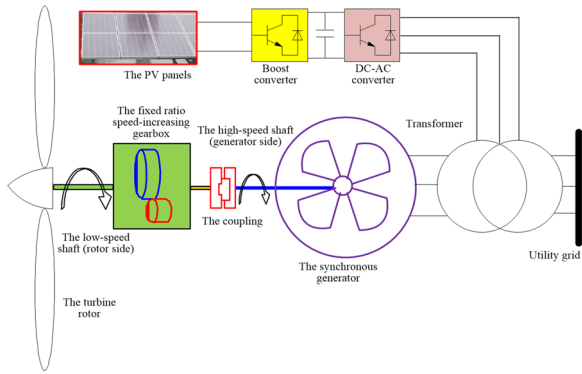


Fig. 1 The integrated offshore wind and photovoltaic power system

2 Integrated offshore wind and photovoltaic power

Figure 1 shows a sketch of the integrated offshore wind and PV power system which consists of an offshore variable speed wind turbine and PV panels that are arranged in an array. All the PV panels are connected to each other and are fixed to the floating wind turbine platform (not shown). By regulating the generator torque, the turbine rotor speed can be continuously varied according to inflow wind conditions, while the generator rotation speed can be kept constant. The PV power system also includes a boost converter connected with the DC link and DC-AC inverter tied to the grid through a transformer. The boost converter acts as a step-up converter and performs maximum power point tracking (MPPT) operation of the PV panels. In order to develop the DRL control strategy to promote the complementarity operation of this offshore wind and PV resources and to reduce energy instability, it is essential to consider the system dynamic response and construct a control-oriented model.

2.1 Dynamic model of the offshore wind turbine

As shown in Fig. 1, the dual-stage mechanical transmission and the electrical generator are important elements for the integrated power system. By varying the generator torque, the rotational speed of the turbine rotor varies, and hence the combined power output of the integrated system can be continuously varied accordingly. When the rotational speed of the wind turbine is changed according to the inflow wind condition to maximize the wind power capture, the rotational speed of the generator is kept constant for connection with the grid. When there is insufficient PV power, the turbine rotor speed and

generator power can be directly regulated such that the overall power generation of the integrated power system will be kept stable.

In this paper, the NREL offshore 5-MW baseline wind turbine is considered for the integrated power system. The wind turbine has a rotor radius of about 63 m, the rated rotor speed is 12.1 rpm, with a rated generator speed of 1173.7 rpm, and gearbox ratio of 97:1. The control of the wind turbine and the integrated power system is focused on region 2 in which the generator torque is controlled to be proportional to the square of the filtered generator speed to maintain a constant (optimal) tip-speed ratio for optimizing power capture and smoothing the overall power output. In this region, the peak power coefficient of 0.482 occurs at a tip-speed ratio of around 8 and a rotor-collective blade-pitch angle of 0.0° .

By neglecting the damping of the mechanical transmission, the dynamic model of the mechanical transmission can be generally formulated as:

$$\begin{cases} J_r \dot{\omega}_r = T_a - T_L \\ J_g \dot{\omega}_g = T_h - T_g \end{cases} \quad (1)$$

where T_a is the low-speed shaft aerodynamic torque, T_g is the high-speed shaft generator torque, T_L and T_h are the respective torques at the gearbox ends of the low-speed and high-speed shafts. ω_r and ω_g are the rotational speeds of the turbine rotor and generator shaft, respectively.

By assuming rigid mechanical transmission and a fixed gear transmission ratio, the model of the transmission can be described as:

$$\begin{cases} T_L \omega_r = T_h \omega_g \\ i_g = \frac{T_h}{T_L} \end{cases} \quad (2)$$

where i_g denotes the transmission ratio of the wind turbine system.

By combining (1) and (2), the dynamic model of the turbine system cast to the low-speed turbine rotor side can be described as:

$$(J_r + J_g i_g^2) \dot{\omega}_r = T_a - T_g i_g \quad (3)$$

where J_r and J_g are the turbine rotor inertia relative to the turbine shaft and generator inertia relative to the high-speed shaft, respectively.

The aerodynamic torque T_a can be described as:

$$T_a = \frac{\pi \rho R^3 v^2(t)}{2\lambda} C_p(\lambda, \beta) \quad (4)$$

where $v(t)$ is the inflow wind speed, R is the turbine radius, ρ is the sea air density, β is the turbine pitch angle which is set to zero in the control region 2 of this work, C_p denotes the power coefficient which indicates the power extraction performance of the turbine, and λ is the tip speed ratio, defined as $\lambda = \omega_r R / v(t)$.

The power coefficient can be defined as [16]:

$$\begin{cases} C_p = c_1 \left(\frac{c_2}{\lambda_i} - c_3 \beta - c_4 \right) e^{\frac{-c_5}{\lambda_i}} + c_6 \lambda \\ \frac{1}{\lambda_i} = \frac{1}{\lambda + 0.08 \beta} - \frac{0.035}{\beta^3 + 1} \end{cases} \quad (5)$$

where the constant coefficients are $c_1 = 0.5176$, $c_2 = 116$, $c_3 = 0.4$, $c_4 = 5$, $c_5 = 21$, $c_6 = 0.0068$.

As shown in Fig. 2, the power coefficient has a clearly nonlinear relationship with respect to the tip speed ratio. There also exists an optimal tip speed ratio such that the

$$\begin{cases} I_{pv} = N_p I_{ph} - N_p I_0 \left[\exp \left(\frac{q(V_{pv} + I_{pv} R_s)}{N_s n k T} \right) - 1 \right] - V_{pv} + (I_{pv} R_s) / R_{sh} \\ I_{ph} = [I_{sc} + K_i (T_k - T_{ref})] * \frac{\lambda_{pv}}{1000} \\ I_0 = I_{pv} \left(\frac{T}{T_{ref}} \right)^3 \exp \left[\frac{q E_{g0}}{n k} \left(\frac{1}{T_{ref}} - \frac{1}{T} \right) \right] \\ I_{rs} = \frac{I_{sc}}{\exp(q V_{oc} / N_s k n T) - 1} \end{cases} \quad (6)$$

power coefficient achieves the maximum value. As shown in Fig. 2, the maximum power coefficient is around 0.48 and the optimal tip speed ratio is around 8.11. Thus, it is clear that the maximum power capture can be achieved if appropriate generator torque control is applied. However, the PV power generation control is also designed by regulating the generator torque such that the overall power generation of the integrated system can be relatively smooth.

By observing (1)–(5), it is clear that the transmission ratio of the wind turbine can be continuously varied by regulating the generator torque according to the turbine rotation speed and inflow wind condition, and hence the maximum power tracking control can be readily achieved

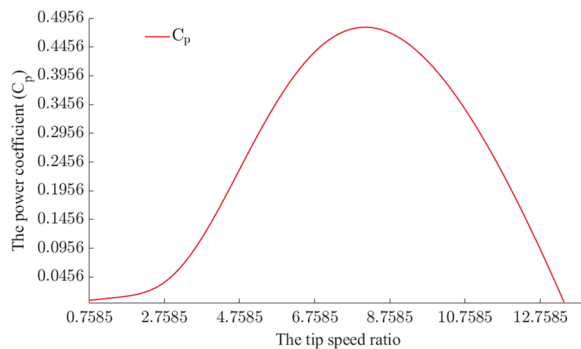


Fig. 2 The wind turbine power coefficient

while the power compensation of the PV is accomplished simultaneously.

2.2 Model of the photovoltaic power

A PV array is composed of multiple series and parallel connected PV modules. Each PV module consists of a number of solar cells connected in series and parallel to obtain the desired output voltage and current. Each solar cell is basically a p-n diode that can convert the incident energy from sunlight directly into electrical energy.

The power generation of the PV cells is dependent on weather conditions including irradiation and temperature. For convenience, the single-diode model is used to describe the PV cell [17], so the relationship between the PV current and voltage can be modelled as:

where

I_{ph} : the light-generated cell photocurrent at the nominal condition of 25°C, 1000 W/m²

R_{sh} : the intrinsic shunt resistance or the equivalent parallel resistance, 500 Ω.

R_s : the equivalent series resistance, 0.1 Ω.

K_i : the short-circuit current/temperature coefficient, 0.0017 A/K.

I_{sc} : the short circuit current, A

T_k : the actual temperature, 298 K.

T_{ref} : the reference temperature, 298 K. λ_{pv} : the irradiation (W/m²), the nominal irradiation is 1200 W/m². q : the electron charge, 1.610⁻¹⁹ C. k : the Boltzmann constant, 1.380510⁻²³ J/K

V_{oc} : the open-circuit voltage, V

N_s : the number of cells connected in series in the given photovoltaic module.

N_p : the number of cells connected in parallel in the given photovoltaic module. n : the ideality factor, 1.6

T : the module operating temperature, 323 K.

T_{ref} : the reference temperature, 298 K.

E_{g0} : the bandgap energy of the semiconductor, 1.1 eV (1 eV = 1.6 × 10⁻¹⁹ J).

3 The DDPG methodology

By considering the control system as an extension to a partially-observable Markov decision process (MDP), an actor-critic architecture DRL algorithm, i.e., deep

deterministic policy gradient (DDPG), is adopted herein to explore and learn the optimal multi-objective control policy. Unlike the existing model-based control algorithms that are heavily reliant on accurate system modeling, the DDPG algorithm is a model-free approach that is well-suited to sequential uncertain optimal control problems.

3.1 The problem formulation

For the integrated power system, the multi-objective value function to be minimized at the time can be formulated as a trade-off between the power generation and power oscillation damping performance, as:

$$\min ER(t) = \alpha_1 \cdot C_p(t) + \alpha_2 \cdot R_p(t) \quad (7)$$

where $ER(t)$ denotes the operational value function, α_1 and α_2 denote the penalty coefficients. $R_p(t)$ denotes the normalized variation rate of the power output of the integrated power system.

The normalized variation rate of the power output of the integrated power system $R_p(t)$ can be defined as:

$$R_p = \left| \frac{\Delta P_{g,pv}}{R_{p,max} \Delta t} \right| \quad (8)$$

where Δt denotes the time interval for evaluating the total power oscillations, $\Delta P_{g,pv}$ denotes the integrated power variation during a time interval, and $R_{p,max}$ denotes the variation rate of the power output of the integrated power system.

As a stochastic control process, MDP is designed to provide the framework for modelling the sequential-decision making problem. MDP can be represented by a tuple of a finite of states, control actions and the state transition probability, and hence can model the interactions between the integrated power system and the DDPG agent.

The main components for MDP can be described as follows:

- (1) *Agent* The DDPG agent aims to generate control actions by gaining experience through repeated interactions with the environment. A well-trained DDPG agent can generate the optimal or near-optimal control policy for the integrated system in real time. The control action is:

$$a(t) \triangleq T_g(t) \quad (9)$$

- (2) *States* As real-time information from the environment, the states are used to indicate the status of the environment. DDPG will make control decisions

based on the obtained state information through interaction with the environment. The state of the integrated system is the turbine rotation speed, as:

$$s(t) \triangleq \{\omega_r(t)\} \quad (10)$$

- (3) *Reward* As the evaluation index of the DDPG agent in MDP, the reward is designed to guide the agent to learn the optimal control policy by minimizing the control objective in (10). The reward is designed as the performance value, as:

$$r(a(t)|s(t)) \triangleq ER(t) \quad (11)$$

- (4) *Constraints* The essential constraints of the integrated system include the constraints for the control actions and states i.e.:

$$\begin{cases} 0 < \omega_r < \omega_{r, \text{rated}} \\ T_{g, \text{min}} < T_g < T_{g, \text{max}} \\ 0 < C_p < 1 \end{cases} \quad (12)$$

where $\omega_{r, \text{rated}}$ denotes the rated turbine rotation speed (1.267 rad/s), $T_{g, \text{min}}$ and $T_{g, \text{max}}$ denote the minimum and maximum generator torques, respectively.

As a discounted sum of the reward function, the accumulated reward $R(t)$ of the exploration process beginning with the state $s(t)$ can be formulated as:

$$\begin{aligned} R(t) &= r(a(t)|s(t)) + \gamma r(a(t+1)|s(t+1)) \\ &\quad + \gamma^2 r(a(t+2)|s(t+2)) + \dots + \gamma^{T-t} r(a(T)|s(T)) \\ &= r(a(t)|s(t)) + \gamma R_{t+1} \end{aligned} \quad (13)$$

where $\gamma \in [0, 1]$ denotes a discount factor, and T denotes the time period of the exploration process.

For the policy π generated by the agent, the value function $Q^\pi(s, a)$ for performing the exploration process can be described as:

$$Q^\pi(s, a) = \mathbb{E}^\pi [R(t)|s = s(t), a = a(t)] \quad (14)$$

The DDPG agent approximates the long-term reward and hence the optimal control policy π^* can be designed as:

$$\pi^* = \arg \max_{\pi} \{Q^\pi(s, a)\} \quad (15)$$

At each discrete time step in MDP, the agent decides a possible action as input for the environment by observing the current state. This results in the next state and a reward from the environment. Consequently, through continuous interactions with the environment and mapping the local states to the control actions at finite time steps, the DDPG agent can derive the optimal control

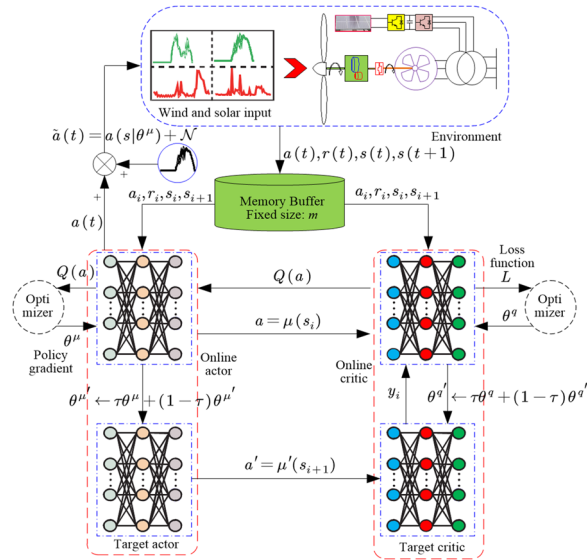


Fig. 3 The DDPG control architecture diagram for the integrated OWT and PV power system

policy π^* such that the maximum accumulative reward over time can be achieved sequentially.

3.2 The implementation procedure

The DDPG agent approximates the long-term reward given actions and observations by using two critic value function representations for different purposes: the critic function provides the judgement of the actor and the actor function learns and updates the control policy by using the minimum value function estimate [18].

As illustrated in Fig. 3, the overall control architecture comprises three parts, i.e., the Actor networks (online actor and target actor networks), the Critic networks (online critic and target critic networks), and the experience replay buffer [19]. The Actor networks are used to map the states to the control action, the Critic networks are employed for estimating the value of state and state-action, while the replay buffer takes charge of storing experiences. The two-copy networks (target actor and target critic networks) are employed to improve the stability of the DDPG algorithm by calculating the target values. In addition, the experience replay buffer is adopted to store a large number of transitions and can randomly sample a mini-batch data from the memory to help break the correlation among training data when updating. The actor network, which is parameterized

by θ^μ , has the output $a(t) = \mu(s(t); \theta^\mu)$, while the output of the target actor network (parameterized by $\theta^{\mu'}$ is $a(t)' = \mu'(s(t+1); \theta^{\mu'})$. The output of the critic network (parameterized by θ^q) is $q = Q(s(t), a(t); \theta^q)$, and the output of the target critic network (parameterized by $\theta^{q'}$) is $q' = Q'(s(t), a(t)'; \theta^{q'})$.

During the MDP training process, the parameters of the actor networks are updated by the gradient descent method in the direction of reducing the loss. Hence:

$$\nabla_{\theta^\mu} J \approx \frac{1}{m} \sum_{i=1}^m [\nabla_a Q(s, a; \theta^q) |s_i, \mu(s_i) \nabla_{\theta^\mu} \mu(s; \theta^\mu) |s_i] \quad (16)$$

where m denotes the batch size.

The parameters θ^q of the critic network are updated by minimizing the following loss function:

$$\mathcal{L} = \frac{1}{m} \sum_{i=1}^m [y_i - Q(s_i, a_i; \theta^q)]^2 \quad (17)$$

$$y_i = r_i + \gamma Q'(s_{i+\Delta t}, \mu'(s_{i+\Delta t}; \theta^{\mu'}); \theta^{q'}) \quad (18)$$

Additionally, a soft update strategy is employed to update the parameters of the target actor and critic networks, as:

$$\begin{cases} \theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'} \\ \theta^{q'} \leftarrow \tau \theta^q + (1 - \tau) \theta^{q'} \end{cases} \quad (19)$$

where $\theta^{\mu'}, \theta^{q'}$ denote the target actor and critic network parameters, respectively. $\tau \in (0, 1]$ denotes the coefficient of the soft update.

The clipped Gaussian noise (or the target policy noise) is also added to the actions to improve the randomness of the control actions, as:

$$\tilde{a}(t) = a(s|\theta^\mu) + \mathcal{N} \quad (20)$$

where $\mathcal{N}$ represents the Gaussian process, which makes the DDPG agent more effectively explore the continuous action domain.

Afterwards, the algorithm starts running with episodic iteration and contains three stages: exploration, learning and convergence. In order that the critic can provide a more accurate judgement and the actor can learn a better control strategy, the actor and critic networks are trained against each other and the main algorithm is designed as the following:

Algorithm: DDPG for the integrated OWT and PV power system

1. Initialize the hyper-parameters θ^μ, θ^q for the actor and critic networks
 2. Initialize the hyper-parameters for the target actor and critic networks $\theta^{\mu'} \leftarrow \theta^\mu, \theta^{q'} \leftarrow \theta^q$
 3. Initialize the m -sized replay buffer $\mathcal{D}$
 4. **For** episode ranging from 1 to M **do**
 5. Reset all the environment for action exploration
 6. Observe the initial state
 7. **for** $t = 1, \dots, T$: **do**
 8. Generate the action based on the current policy and sampled noise
 9. Implement the action and obtain the reward and the next state
 10. Store the transition $a(t), r(t), s(t), s(t+1)$ in the replay buffer $\mathcal{D}$
 11. **if** $\mathcal{D}$ is full **then**
 12. Randomly sample m -sized transitions
 13. **if** episode is even **then**
 14. Update θ^μ with Eq. (16)
 15. Update θ^q with Eq. (17), Eq. (18)
 16. Update $\theta^{\mu'}, \theta^{q'}$ with Eq. (19)
 17. **end if**
 18. **end if**
 19. **end for**
 20. **end for**
-

4 Results and discussions

The potential and effectiveness of the DDPG control for the integrated OWT and offshore PV power system are evaluated based on design experiments. The performance of the optimal power generation and power oscillation suppression is also verified. The validations are conducted based on comparison with the results obtained by using a PI (Proportional-Integral) type conventional controller widely used in the wind energy industry.

4.1 The design experiments

The employed OWT is a three-bladed 5 MW offshore upwind horizontal-axis variable speed wind turbine with a monopile type platform. The power coefficient is a function of the tip-speed ratio and blade-pitch angle [20], while the peak power coefficient is 0.48 at the tip speed ratio of 8.116. The minimum generator speed is 670 rpm which corresponds to the minimum turbine rotation speed of 6.9 rpm. The PV cells are designed and grouped in larger units called PV modules, which are further interconnected in a series-parallel configuration to form a PV array.

The main parameters for the case study are provided in Table 1.

The OWT is initially operated under below-rated wind conditions (7–10 m/s) and the PV panels operate based on the maximum power point tracking mode. The DDPG algorithm and MDP are implemented by using the Tensorflow in Python. The MDP model is designed with a discrete time step of 0.1 min and the simulation time period is 6 min. All the DDPG parameters have been carefully tuned to obtain satisfactory

performance of the integrated power system. Both the actor and critic networks have five fully-connected layers with 50 neurons adopted by rectified linear units to each layer. They are then fed into the fully-connected layer of the actor and critic networks to approximate the action and Q value, respectively. The dropout layers are also employed to avoid the vanishing gradient problem. The DDPG algorithm are trained for 300 episodes to learn the optimal generator torque control strategy.

As shown in Fig. 4, the turbulent inflow wind speed for the OWT is generated with the inflow wind speed to be within the range of (7, 10) m/s, and is employed for the optimal peak power capture and power smoothing of the integrated power system.

As shown in Figs. 5 and 6, 100 PV modules are connected together to form a PV array for validating the power generation control of the integrated OWT and PV power. For each PV module, the PV current varies within 30–320 A and the PV voltage varies within 1–48 V, while the PV power output varies around 0.5 MW. The typical PV array characteristics are employed as the exogenous power input for testing the DDPG method since the PV power acts as external perturbation for the control design.

Table 1 Main design parameters for integrated power system

Parameters	Values
Turbine blade radius R	63 m
Wind turbine rotor inertia J_r	38,759,227 kg m ²
Generator inertia J_g	534.2 kg m ²
Gearbox ratio i_g	97:1
Inflow wind speed range	7–10 m/s
Generator torque range	1.79–39.9 kNm
Turbine rotation speed range	0.1–1.267 rad/s
Maximum power variation ratio of the integrated system	4.535 MW/s
Minimum and maximum PV power output	0.4 MW, 0.58 MW
Minimum and maximum integrated power output	0.575 MW, 5.11 MW
Number of cells connected in parallel, N_p	100
Number of cells connected in parallel, N_s	10
Short circuit current, I_{sc}	10 A
Open-circuit voltage, V_{oc}	28 V

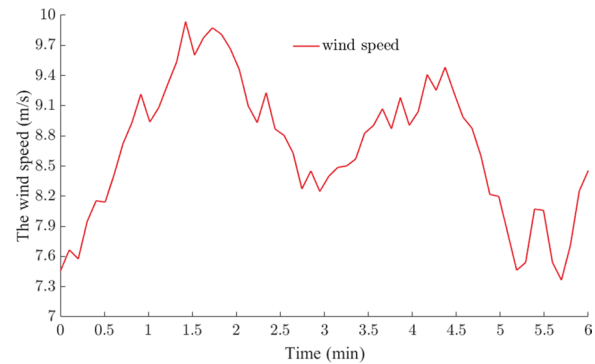


Fig. 4 The inflow wind speed for the wind turbine

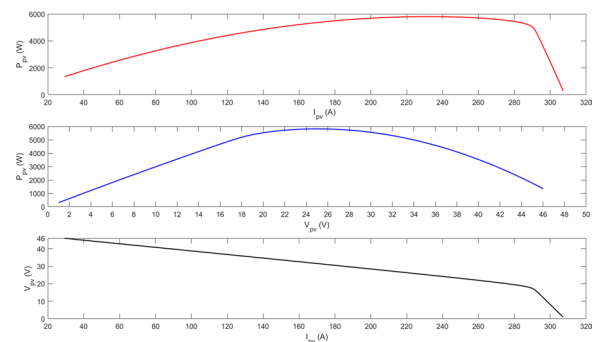


Fig. 5 The PV panel characteristics

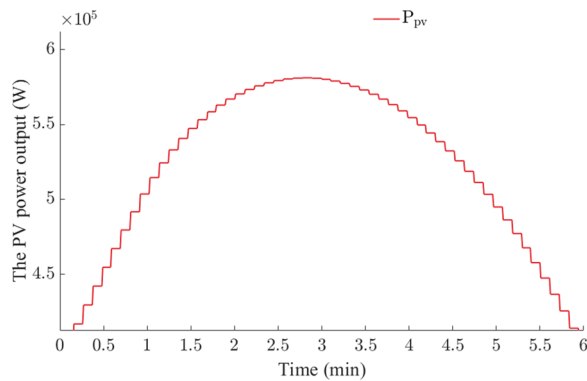


Fig. 6 The power generation of the PV panel used as input for the control design

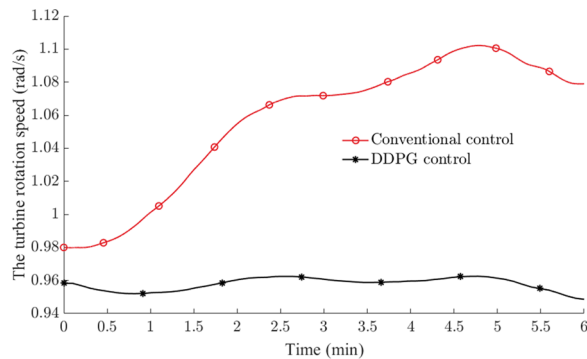


Fig. 7 The turbine rotation speed variations

4.2 The power generation performance

As shown in Fig. 7, the turbine rotor speed can be well regulated such that the maximum active power from the wind can be extracted when the DDPG control is employed while the rotation speed cannot be well regulated to track the optimal value when the conventional control is employed. The result suggests a promising potential in the use of the DDPG in achieving smooth wind power generation and regulation, which is particularly important for electric power with high penetration of intermittent renewables.

As shown in Fig. 8, the integrated power system with the DDPG control can quickly respond to sudden changes of the inflow wind condition. Because of the sufficient capability, the power coefficient of the OWT can be maintained around the optimal value using the conventional control while the power coefficient will vary such that the overall power variations can be compensated for when the DDPG control is used. The result indicates that it is possible to adaptively regulate the power coefficient by using the DDPG control as compared with the conventional control when considering more control objectives.

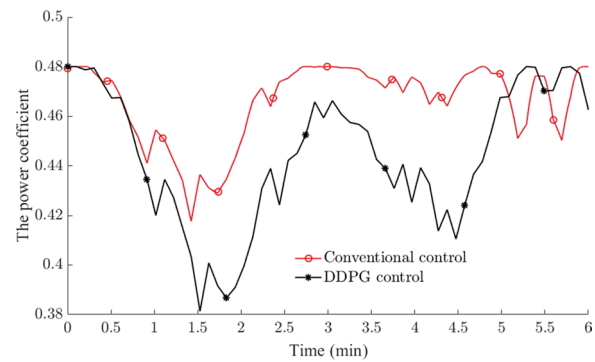


Fig. 8 The power coefficient

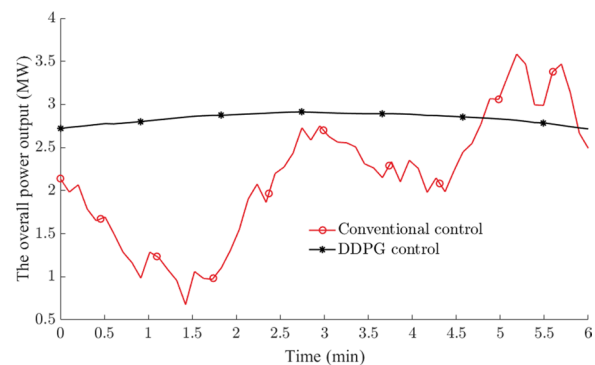


Fig. 9 The offshore OWT and PV power output

As shown in Fig. 9, the total power generated from the OWT and PV array has significant oscillations because of the intermittent wind inflow and the PV power variations when using the conventional control method. The rapid variations in the generated power will in turn cause mechanical vibrations on the OWT blades and tower. By using the DDPG control for the power oscillation damping, it is clear that variations of the integrated power generation can be reduced, making the OWT complement the PV power outputs and smoothing the total power fluctuations.

4.3 The training performance of DDPG

As shown in Fig. 10, the episode score converges quickly to around 0.45 at about 50 episodes for the DDPG training despite the initial oscillations of the episode score values. Therefore, the DDPG method has good convergence performance and it is possible to design and apply a well-trained DDPG agent in the control of the integrated power system.

As illustrated in Fig. 11, the rolling score for DDPG also converges eventually to around the value of around 0.45 which is the same as the episode score. In practice, the rolling score is a filtered version of the episode score and

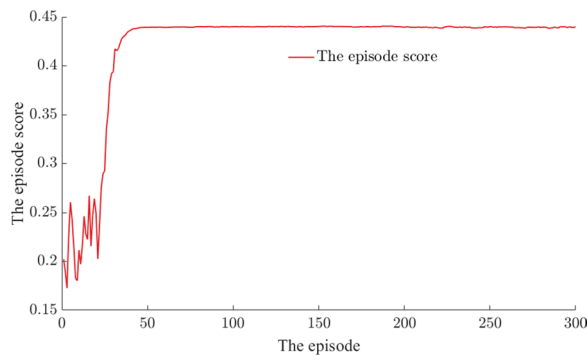


Fig. 10 The episode score variations

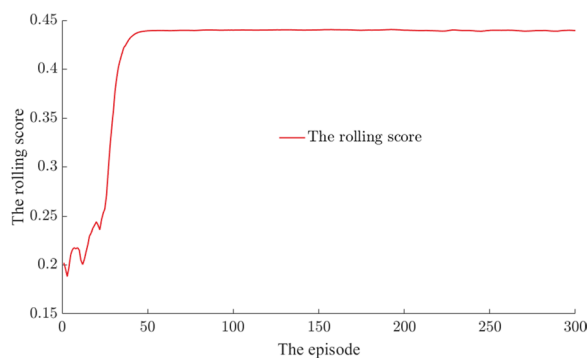


Fig. 11 The rolling score variations

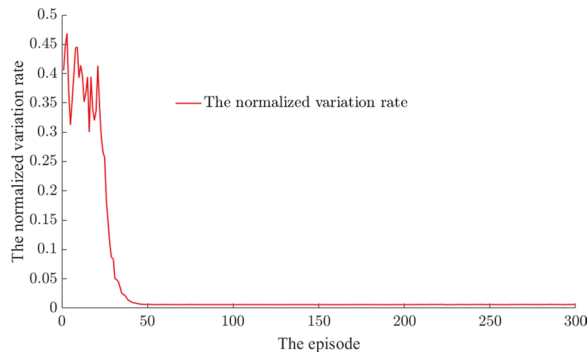


Fig. 12 The normalized variation rate

it is clear that the DDPG control method will converge to the optimal rolling score value. The results clearly reveal that it is promising to apply the DDPG control in the integrated power system such that the optimal rolling score can be achieved.

As observed from Fig. 12, the normalized variation rate of the power output of the integrated power system per episode decreases and reaches a steady state value of around 0 quickly, within about 50 episodes, indicating that the DDPG

control of the integrated power system is closed loop stable and can well achieve the joint operation of power efficiency maximization and power oscillation damping.

As illustrated in the above validation results, it is clear that the DDPG method has relatively good performance and the following advantages:

- (a) The proposed DDPG method has a faster response than the conventional control method. By using the well-trained DDPG model, it is possible to generate the control action within a very short time interval and hence the trained agent in DDPG will be actively involved in the control of the integrated power system.
- (b) The DDPG method can provide a feasible solution to the constrained finite time optimal multi-objective control problem. Hence, by using the DDPG method, the control objectives of improving energy efficiency and smoothing power oscillation of integrated offshore wind and photovoltaic power can be both taken into consideration.
- (c) The proposed RL method is incorporated with a feasible or allowable range of the control actions such that the safety of the integrated power system can be checked in a timely manner in case the agent generates wrong actions and causes failed future power generation. Therefore, the proposed DDPG method is safer than the conventional control algorithm.
- (d) The proposed DDPG method is model-free and does not require prior domain knowledge or a pre-defined rule to decide how to choose an action.

Considering the above advantages of the proposed method, it is promising to use the DDPG control method in real-world applications. Therefore, in order to implement the DDPG method, the hyperparameters of the DDPG model can be adaptively tuned to match the real-world control plant. Also, the valuable knowledge from the control plant can be collected before the implementation such that the decision-making problems can be quickly addressed. In addition, a laboratory-level testbed using the Raspberry Pi boards [21] could be first constructed by specifying key technical details and implementation procedure over which the DDPG algorithm can be tested and validated such that appropriate control actions can be generated before practical application.

5 Conclusion

The paper has presented the design, dynamics and DDPG optimal control for an integrated OWT and PV power system, considering the damping of power oscillations in particular. A variable-speed OWT with electrical

generator torque control has been used in the integrated offshore system for active power regulation and oscillation damping. The results from design experiments indicate that the OWT responds quickly to maximize the total power generation when using the DDPG control, while the power oscillations can also be better damped by regulating the generator torque when using the DDPG control. The results indicate the potential for synergies between offshore wind and PV power utilization, and it is clear that the complementary operations of the integrated system can achieve peak power shaving performances. In future work, a scaled-down prototype of the offshore WT and PV station will be built and experimental verifications of the proposed control approach will be conducted.

Acknowledgements

The work was supported by "the Fundamental Research Funds for the Central Universities. No. 2042022gf0008. The work was supported by the Guangxi Science and Technology Base and Talent Special Project No. 2019AC20266. The work was supported by the Natural Science Foundation of Guangxi under Grant No. 2019JJB160062. The work was supported by Guangdong Basic and Applied Basic Research Foundation (Grant No. 2019A1515110709).

Author contributions

XY: Conceptualization, Methodology, Writing—original draft, Data curation, Visualization, Investigation, Validation. Both authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

Data and materials will be available upon reasonable request.

Declarations

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received: 4 July 2022 Accepted: 15 May 2023

Published online: 31 May 2023

References

- Parastegari, M., Hooshmand, R. A., Khodabakhshian, A., et al. (2015). Joint operation of wind farm, photovoltaic, pump-storage and energy storage devices in energy and reserve markets. *International Journal of Electrical Power & Energy Systems*, 64, 275–284.
- Wu, Y., & Zhang, T. (2021). Risk assessment of offshore wave-wind-solar-compressed air energy storage power plant through fuzzy comprehensive evaluation model. *Energy*, 223, 120057.
- Wang, X., Mei, Y., Kong, Y., et al. (2017). Improved multi-objective model and analysis of the coordinated operation of a hydro-wind-photovoltaic system. *Energy*, 134, 813–839.
- Golroodbari, S. Z. M., Vaartjes, D. F., Meit, J. B. L., et al. (2021). Pooling the cable: A techno-economic feasibility study of integrating offshore floating photovoltaic solar technology within an offshore wind park. *Solar Energy*, 219, 65–74.
- Tang, Y., Fang, G., Tan, Q., et al. (2020). Optimizing the sizes of wind and photovoltaic power plants integrated into a hydropower station based on power output complementarity. *Energy Conversion and Management*, 206, 112465.
- Miller, I., Gençer, E., & O'Sullivan, F. M. (2018). A general model for estimating emissions from integrated power generation and energy storage. Case study: Integration of solar photovoltaic power and wind power with batteries. *Processes*, 6(12), 267.
- López, M., Rodríguez, N., & Iglesias, G. (2020). Combined floating offshore wind and solar PV. *Journal of Marine Science and Engineering*, 8(8), 576.
- Wu, Y., Zhang, T., Xu, C., et al. (2019). Optimal location selection for offshore wind-PV-seawater pumped storage power plant using a hybrid MCDM approach: A two-stage framework. *Energy Conversion and Management*, 199, 112066.
- Elgammal, A., & Jagessar, M. (2020). Optimal control strategy for a marine current farm integrated with a hybrid PV system/offshore wind/battery energy storage system. *European Journal of Electrical Engineering and Computer Science*, 4(4).
- Lund, H. (2006). Large-scale integration of optimal combinations of PV, wind and wave power into the electricity supply. *Renewable Energy*, 31(4), 503–515.
- Lu, T., Sherman, P., Chen, X., et al. (2020). India's potential for integrating solar and on-and offshore wind power into its energy system. *Nature communications*, 11(1), 1–10.
- Feng, L., Zhang, J., Li, G., et al. (2016). Cost reduction of a hybrid energy storage system considering correlation between wind and PV power. *Protection and Control of Modern Power Systems*, 1(1), 1–9.
- Anjaiah, K., Dash, P. K., & Sahani, M. (2022). A new protection scheme for PV-wind based DC-ring microgrid by using modified multifractal detrended fluctuation analysis. *Protection and Control of Modern Power Systems*, 7(1), 1–24.
- El-Kareem, A., Hesham, A., Abd Elhameed, M., et al. (2021). Effective damping of local low frequency oscillations in power systems integrated with bulk PV generation. *Protection and Control of Modern Power Systems*, 6(1), 1–13.
- Karanam, A. N., & Shaw, B. (2022). A new two-degree of freedom combined PID controller for automatic generation control of a wind integrated interconnected power system. *Protection and Control of Modern Power Systems*, 7(1), 1–16.
- Xia, Y., Ahmed, K. H., & Williams, B. W. (2012). Wind turbine power coefficient analysis of a new maximum power point tracking technique. *IEEE Transactions on Industrial Electronics*, 60(3), 1122–1132.
- Pandiarajan, N., Ramaprabha, R., & Muthu, R. (2012). Application of circuit model for photovoltaic energy conversion system. *International Journal of Photoenergy*, 2012.
- Rui, X., Su, R., Wu, X., et al. (2014). The conceptual design of grid-connected wind turbine based on speed regulating differential mechanism. *Journal of Mechanical Science and Technology*, 28(6), 2215–2220.
- Liu, T., Hu, X., Hu, W., et al. (2019). A heuristic planning reinforcement learning-based energy management for power-split plug-in hybrid electric vehicles. *IEEE Transactions on Industrial Informatics*, 15(12), 6436–6445.
- Kühne, P., Pöschke, F., & Schulte, H. (2018). Fault estimation and fault-tolerant control of the FAST NREL 5-MW reference wind turbine using a proportional multi-integral observer. *International Journal of Adaptive Control and Signal Processing*, 32(4), 568–585.
- Zhang, X., Lu, R., Jiang, J., et al. (2021). Testbed implementation of reinforcement learning-based demand response energy management system. *Applied Energy*, 297, 117131.