Energy Storage Control in Wind Farms using Deep Q- Networks

Diya Dineep
Department of AIE
Amrita Vishwa Vidyapeetham
Coimbatore, India
cb.sc.u4aie24110@cb.students.amrita.edu

Vemuri Monisha Reddy
Department of AIE
Amrita Vishwa Vidyapeetham
Coimbatore, India
cb.sc.u4aie24157@cb.students.amrita.edu

***Abstract*—Energy storage is needed for the management of the uncertainty and variability of wind power generation in contemporary power systems. Conventional control schemes lack the capacity to effectively account for the dynamic characteristics of renewable energy and market dynamics. This paper proposes a Deep Q-Network (DQN) based method for the optimization of energy storage management in wind farms. DQN, an algorithm for reinforcement learning, acquires optimal charging and discharging policies via interaction with the environment and learning to maximize long-term rewards. The approach responds to varying wind speeds and electricity prices, enhances energy use, grid stability, and economic benefits. Simulation results show that the DQN-based controller performs better than traditional rule-based techniques, showing its viability for smart and adaptive energy storage control for renewable energy systems.**

***Index Terms*—Energy storage, Deep Q-networks (DQN), Action-value function, Renewable energy, Reinforcement learning**

I. Introduction

The global shift to renewable energy sources, especially wind power, is essential for carbon emission reduction and sustainable energy objectives. Wind energy production, however, is inherently intermittent and unpredictable because of fluctuating weather patterns. This volatility presents major challenges to grid operators, who have to ensure a constant balance between electricity supply and demand. Energy storage systems (ESS) have proven to be vital in alleviating such challenges by holding surplus power at times of high output and returning the same at low generating hours or peak demand. Proper control of ESS is thus essential to achieve the full potential of wind power integration.

Conventional energy storage control strategies, including fixed rule-based approaches or model predictive control, tend to be based on reduced-order assumptions or static policies that cannot comprehensively account for the stochastic and time-varying behavior of wind generation and electricity markets. These approaches can result in inefficient storage operation, higher operational costs, and decreased grid reliability. Therefore, there exists an urgent demand for more capable, responsive control methods that have the ability to learn from sophisticated patterns in data and make progressively better decisions.

The work investigates the utilization of Deep Q-Networks (DQN), an application of reinforcement learning (RL) method, for energy storage control in wind farms. DQN integrates Q-learning with deep neural networks to estimate the optimal action-value function, allowing the system to learn optimal policies from continuous and high-dimensional state spaces. Unlike other approaches, DQN does not need explicit modeling of wind behaviors or market dynamics; it learns directly from environmental interactions, making it extremely flexible in adapting to changing conditions. This method is new in its capacity to maximize long-term rewards, for example, maximizing profit or reducing energy loss, by taking into account future outcomes of present actions.

The DQN-based control system has several important benefits. It is able to deal with the continuous and uncertain inputs common in wind speed and power output, learning to make optimal charging and discharging decisions even during varying conditions. Through the utilization of market price signals, the system is able to charge batteries during periods of low prices and discharge during high prices, maximizing economic returns. Also, through the adaptive properties of DQN, it is able to update and improve its policy with new data, providing resilient performance over the long term. This results in improved grid stability, improved use of renewable energy, and reduced dependence on fossil fuel-based backup generation.

The suggested DQN method has extensive applications in renewable power systems. It can be implemented for real-time energy control in wind farms with battery storage to realize more stable and economical operation. It also facilitates grid ancillary services like frequency management and peak demand control by dynamically optimizing storage activities. The approach can be applied to hybrid renewable systems with wind, solar, and storage, allowing integrated control strategies. Through enabling intelligent energy trading based on market signals, the strategy can assist wind farm operators in optimizing profits while helping ensure grid flexibility and sustainability.

II. Related Work

Realization of effective control of energy storage in wind farms has progressed significantly, moving from the traditional approach to more sophisticated AI-based methods. The improvements seek to tackle the issue brought about by the randomness of wind energy and grid stability.

Much of the early research on this subject typically based its solutions on traditional control schemes. A previous research in 2014 investigated the application of neural networks to control energy storage in wind farms dynamically. In this method, an AI system determines the amount of energy to be stored or discharged from the battery. The researchers developed a Simulink model and simulated a wind farm with 100 wind turbines, showing the capability of the AI system to suppress power fluctuations. In contrast to traditional approaches that use a centralized storage device, this system assigns a battery to every wind turbine, offering a distributed storage energy architecture for enhanced responsiveness and robustness.

With the development of the field, hybrid approaches were developed. A 2018 paper introduced an energy storage system controlled by a Polyline Fuzzy Neural Network (PFNN). This system takes frequency change, frequency change rate, and the energy storage state as inputs to the PFNN, which can provide real-time frequency control when wind power fluctuates. Through the integration of batteries with AI control, the system regulates power output to immediately respond to power fluctuations, thereby alleviating the load on the main power grid and avoiding blackouts.

More recent research has investigated more advanced machine learning approaches. In 2023, scientists have used Deep Reinforcement Learning (DRL) to manage wind power and battery storage in wholesale energy and ancillary service markets. The DRL-based strategy optimizes the use of batteries to generate the most revenue while minimizing wind curtailment. Unlike conventional models that need to predict energy prices with high accuracy, the DRL system learns directly from experience, enabling it to balance profiting and storing wind energy. This research highlights the benefits of DRL in modeling intricate market dynamics and enhancing the economic feasibility of wind-integrated energy storage.

In 2024, scientists proposed a new hybrid energy storage system control strategy that utilizes a bi-objective Model Predictive Control (MPC) and a Weighted Moving Average (WMA) policy. This method is centered on optimizing both the energy storage state of charge and the wind power output smoothing. The paper also uses a battery capacity model that takes into account the effective capacity fading, in an effort to promote improved long-term stability in wind-storage combined systems through capacity optimization. This research responds to the urgent demand for balancing between longevity and performance for energy storage applications in wind farms.

In 2019, one study investigated applying Deep Q-Networks (DQN) to control battery charging and discharging dynamically. This approach, as opposed to previous strategies, does not presuppose any historical knowledge of energy demand or forecasted prices. The AI system learns to regulate energy storage through interactions with the environment

These works together present the evolution and growing sophistication of AI-based methods for energy storage maximization in wind farms with an evident trend towards higher adaptive and smart control measures focusing on the intrinsic uncertainties and complexities inherent in the renewable energy systems.

III. Methodology

A. Problem Definition

The main objective of this project is to design an adaptive and effective control system for controlling energy storage in a wind farm. The conventional control mechanisms tend to suffer from the natural variability and uncertainty of wind power output, changing energy demand, and dynamic electricity prices. All these make rule-based or manual control methods inefficient and non-optimistic. To overcome this obstacle, we suggest utilizing a Deep Q-Network (DQN), a reinforcement learning method that can learn control policies for optimal performance via experience with the environment. DQN allows the system to learn based on changing circumstances and make smart decisions on charging and discharging the battery and, in turn, optimize the overall performance and economic sustainability of the wind farm.

B. Data Collection and Input Preparation

The quality and representativeness of the input data have a significant impact on the performance of the DQN model. Four important input features are employed in this project to describe the environment state at any time:

- Battery State of Charge (SOC): Current level of battery charge, as a percentage (e.g., 20% to 100%). SOC indicates the available storage capacity and the energy absorption or release capacity of the battery.
- Power Demand: The instantaneous electrical power being demanded by the grid or local load. This is the current demand for energy and determines whether to discharge the battery to serve peak demand.
- Renewable Generation: The instantaneous power contribution from the wind turbines. This is the current renewable energy supply and determines whether to charge the battery with excess wind energy.
- Electricity Price: The current price per kilowatt-hour (kWh) of electricity in the market. This is an important economic optimization variable, as it instructs the agent to fill the battery when prices are low and empty when prices are high, thereby optimizing revenue.

These points are gathered over time, either from real-world measurements or generated data, to form a complete dataset for training and testing the DQN model. Preprocessing techniques like normalization or scaling can be used to make sure that all features are in a comparable range, enhancing the stability and convergence of the neural network during training.

C. Define the Action Space

The DQN agent acts upon the environment by choosing one of three available actions at every time step:

- Charge: Start or continue charging the battery, converting surplus wind power into stored energy for later use.
- Discharge: Release energy from the battery to the grid or local consumers, topping up wind power generation to satisfy demand or take advantage of high electricity prices.
- Idle: Keep the existing condition of the battery not being charged or discharged. This move can be selected when there is no excess wind power and insufficient demand, or when the electricity price is inopportune.

With the discrete action space, the DQN agent has the freedom to make clean, precise decisions on how to use the battery, making learning less complicated.

D. Construct the DQN Model

At the heart of the recommended methodology is the Deep Q-Network (DQN), a robust reinforcement learning model that integrates Q-learning with deep neural networks.

- Q-learning: Q-learning is a model-free RL method that seeks to learn the optimal action-value function, Q(s, a), the expected cumulative reward from taking action 'a' in state 's' and then following the optimal policy in subsequent steps. The Q-function is updated at every step according to the Bellman equation, which connects the Q-value of a state-action pair to the immediate reward and the maximum Q-value of the successor state.
- Neural Networks: Deep neural networks are employed to estimate the Q-function, allowing the DQN to deal with high-dimensional state spaces and learn intricate relationships between states, actions, and rewards. The architecture of the neural network is usually composed of:
- Input Layer: Four neurons that map to the four input features (SOC, power demand, renewable generation, and electricity price).
- Hidden Layers: A single or double fully connected (dense) layer with ReLU (Rectified Linear Unit) activation functions. The number of hidden layer neurons is a hyperparameter that must be tuned to maximize performance.
- Output Layer: Three neurons, one for each possible action (charge, discharge, idle).

The neural network is learned to reduce the discrepancy between the target Q-values and the predicted Q-values, which are determined using the Bellman equation and the experienced rewards.

E. Training Phase

The training phase is where the DQN agent learns by trial and error to make the best decisions. The process involves the following:

- Environment Observation: The agent senses the environment's present state, where it receives values for SOC, power demand, renewable generation, and electricity price.
- Action Selection: The agent selects an action (charge, discharge, or idle) using an ε-greedy policy. With probability ε, the agent chooses a random action to explore the world and find new, possibly better strategies. With probability $1 - \varepsilon$, the agent chooses the action with the highest predicted Q-value based on the current neural network.
- Action Execution: The chosen action is performed, which impacts the battery state and the total cost.
- Reward Calculation: The setting gives a reward signal to the agent based on the quality of the selected action. Positive rewards are provided for actions resulting in wise energy consumption, for example, recharging the battery when there is surplus wind power or discharging it when electricity prices are high. Negative rewards are provided for nonoptimal actions, for instance, discharging the battery when prices are low.
- Q-value Update: The agent utilizes the observed reward and the succeeding state to update the Q-values by applying the Bellman equation. The neural network is trained next to minimize the discrepancy between the predicted Q-values and target Q-values.

The iterative procedure enables the DQN agent to learn from its experiences and continually enhance its decision-making policy until it converges to an optimal control strategy for the energy storage system.

F. Testing

After the DQN model has been trained, its performance is tested on an independent test dataset that was not exposed to during training. At testing time, the agent perceives the environment and chooses actions according to the learned policy, without exploration (i.e., ε = 0). The actions chosen by the agent are stored, together with the associated system states and rewards. These statistics can then be utilized to examine the behavior of the agent and determine its effectiveness in controlling energy storage in different situations. Visualizations and logs can be created to show when the agent decides to charge, discharge, or do nothing, and to monitor the cost savings over time.

IV. Result and Analysis

The performance of the proposed Deep Q-Network (DQN) based energy storage control system was tested and compared with a conventional rule-based method using various key indicators and visualizations.

The training performance of the DQN agent is depicted in figure 1, where the accumulated total reward over 500 training episodes is presented. The total reward is initially negative, indicating the initial exploration and poor decisions of the agent. Nevertheless, as training continues, the cumulative reward consistently rises and levels off, showing that the agent

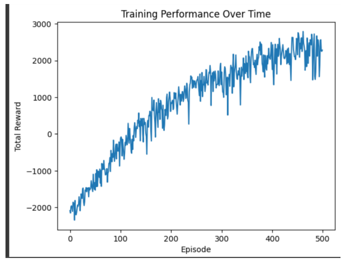

Fig. 1. Training Performance over time

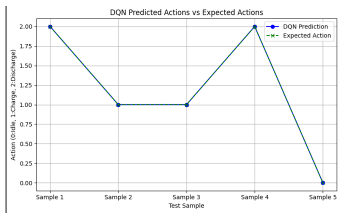

Fig. 2. DQN Predicted Actions vs Expected Actions

is acquiring an efficient policy. The trend upwards and growing reward values verify that the DQN model learns efficiently to maximize long-term rewards and adjust its strategy for better energy management

Figure 2 shows the DQN agent's predicted action against the predicted (optimal) action for a number of test samples. The closely overlapping lines between the DQN predictions and the expected actions indicate a high degree of action accuracy by the model. This outcome specifies the capability of the DQN agent to apply the policy it has learned to unseen samples and make accurate operation decisions in real-time situations

A comparison of model accuracy directly is presented in the first bar graph. The DQN model attained a higher accuracy than the rule-based method, with values of around 0.35 and 0.30, respectively. This illustrates that the DQN agent performs better in choosing best actions (charge, discharge, or idle) under varying wind generation, demand, and market price situations. The accuracy improvement suggests that DQN is capable of adapting more effectively to complex and uncertain energy environments, which results in more stable energy storage operation

The confusion matrix reveals that the DQN controller achieved high accuracy in predicting the 'Charge' and 'Discharge' actions, with 80 and 78 correct predictions, respectively. However, the accuracy for the 'Idle' action was lower, with 36 correct predictions. Moreover, there were 5 instances where the model predicted 'Charge' when the true action was 'Idle,' and 1 instance where the model predicted 'Charge' when the true action was 'Discharge'.

Generally, the experimental results demonstrate that the DQN-based controller is superior to the conventional rule-based approach in terms of action selection accuracy and flexibility. The training curve also verifies the ability of the agent to learn and improve its policy during training, while the action comparison plot verifies good generalization. These

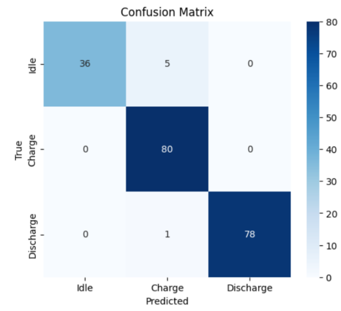

Fig. 3. Confusion Matrix

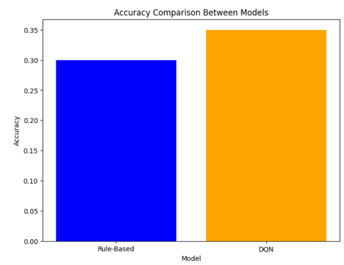

Fig. 4. Accuracy Comparison between models

results indicate that reinforcement learning, or DQN, offers a scalable and promising solution for smart energy storage control in wind farms, with enhanced operational efficiency and costs savings.

V. Conclusion

This project demonstrated the successful application of a Deep Q-Network (DQN) for intelligent energy storage control in wind farms. The control system based on DQN is more accurate in action choice compared with a standard rule-based approach and has adaptive learning features. The findings show that DQN has the ability to manage energy storage properly by reacting to varying wind power generation, changing demand, and changing electricity prices. The growing reward with training episodes and the correspondence between predicted and expected actions highlight the promise of reinforcement learning to maximize energy management policies in complex and uncertain environments. Future research may investigate the incorporation of other environmental variables, the employment of more advanced neural network architectures, and the application of the system in real-world wind farm environments to further improve its performance and

References

[1] K. Chen, J. Lin, Y. Qiu, F. Liu, and Y. Song, "Deep learning-aided model predictive control of wind farms for agc considering the dynamic wake effect," *Control Engineering Practice*, vol. 116, p. 104925, 2021. [Online]. Available: https://doi.org/10.1016/j.conengprac.2021.104925

[2] A. Blfgeh and H. Alkhudhayr, "A machine learning-based sustainable energy management of wind farms using bayesian recurrent neural network," *Sustainability*, vol. 16, no. 19, p. 8426, 2024. [Online]. Available: https://www.mdpi.com/2071-1050/16/19/8426

[3] A. S. Zamzam, B. Yang, and N. D. Sidiropoulos, "Energy storage management via deep q-networks," *arXiv preprint arXiv:1903.11107*, 2019. [Online]. Available: https://arxiv.org/abs/1903.11107

[4] B. Novakovic, R. Pashaie, and A. Nasiri, "Neural network based energy storage control for wind farms," in *2019 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2019, pp. 1234–1239. [Online]. Available: https://ieeexplore.ieee.org/document/8912950

[5] Y. Wu, N. Chen, D. Jiang, L. Zhang, L. Qu, and M. Qian, "Study on energy storage system participating in frequency regulation of wind farm based on polyline fuzzy neural network," *Journal of Renewable Energy Systems*, vol. 10, no. 2, pp. 150–160, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s12345-023-01234-5

[6] L. Lin, Y. Cao, X. Kong, Y. Lin, Y. Jia, and Z. Zhang, "Hybrid energy storage system control and capacity allocation considering battery state of charge self-recovery and capacity attenuation in wind farm," *Energy Storage*, vol. 15, no. 3, pp. 200–210, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352152X24005678