Reinforcement-Learning-Based Virtual Energy Storage System Operation Strategy for Wind Power Forecast Uncertainty Management

Eunsung Oh

Department of Electrical and Electronic Engineering, Hanseo University, Chungcheongnam-do 31962, Korea; esoh@hanseo.ac.kr

Received: 31 August 2020; Accepted: 11 September 2020; Published: 15 September 2020

Abstract: Uncertainties related to wind power generation (WPG) restrict its usage. Energy storage systems (ESSs) are key elements employed in managing this uncertainty. This study proposes a reinforcement learning (RL)-based virtual ESS (VESS) operation strategy for WPG forecast uncertainty management. The VESS logically shares a physical ESS to multiple units, while VESS operation reduces the cost barrier of the ESS. In this study, the VESS operation model is suggested considering not only its own operation but also the operation of other units, and the VESS operation problem is formulated as a decision-making problem. To solve this problem, a policy-learning strategy is proposed based on an expected state-action-reward-state-action (SARSA) approach that is robust to variations in uncertainty. Moreover, multi-dimensional clustering is performed according to the WPG forecast data of multiple units to enhance performance. Simulation results using real datasets recorded by the National Renewable Energy Laboratory project of U.S. demonstrate that the proposed strategy provides a near-optimal performance with a less than 2%-point gap with the optimal solution. In addition, the performance of the VESS operation is enhanced by multi-user diversity gain in comparison with individual ESS operation.

Keywords: energy sharing; energy storage system; reinforcement learning; renewable; uncertainty; virtual energy storage system; wind power

1. Introduction

The use and development of renewables has grown continuously in the power sector owing to climate change, with renewable power generation becoming second in the electricity mix in 2018 [1]. Renewables installed more than 200 gigawatts in 2019, which is the largest increase to date [2]. It is expected that renewable-based power capacity will grow by 50% between 2019 and 2024 [3]. In particular, forecasts predict that wind capacity installations will triple by 2024 [3].

Wind power generation (WPG) is subject to high fluctuations and intermittent properties. The characteristics of WPG make it difficult to ensure power system reliability [4,5]. Although various wind power forecasting methods such as the ensemble method [6], aggregated probabilistic method [7], and machine learning-based method [8] have been researched, uncertainty cannot be completely eliminated owing to the nature of wind-resource phenomena.

An energy storage system (ESS) plays an essential role in managing the uncertainty of WPG [9]. ESSs for WPG are used in various applications such as frequency regulation [10], ramp rate mitigation [11], and demand response [12]. The basic role of an ESS is to charge the surplus energy and discharge the stored energy according to the operational objective. Therefore, the primary issue of the usage of ESS is the effective operation of the ESS. Gomes et al. proposed a stochastic mixed-integer linear programming approach to manage the mismatching of the renewable power generation uncertainty...
using ESS operation in the day-ahead market [13]. Sperstad and Korpas presented stochastic ESS scheduling over an extended planning horizon [14]. Kalavani et al. proposed stochastic ESS scheduling, considering a demand response [15]. Yuan et al. suggested a revised genetic algorithm (GA) to solve the ESS operation optimization problem for economic dispatch [16]. Khare proposed particle swarm optimization and chaotic particle swarm optimization algorithms to minimize energy cost in renewable systems with ESS [17]. Liu et al. suggested a chicken swarm optimization algorithm to improve the reliability of ESS for renewables [18]. Oh and Son presented a frequency domain analysis-based ESS operation algorithm to reduce the uncertainty of WPG [19]. These studies show that ESSs effectively manage uncertainty to enhance the utilization of WPG. However, there are two constraints: First, the performance of the ESS is strongly related to the ESS capacity [19–21]. Second, for effective ESS operation, environment modeling of the characteristics of WPG is required in conventional stochastic and meta-heuristic approaches.

This study focuses on reinforcement learning (RL)-based virtual ESS (VESS) operation to manage the uncertainty of WPG. VESS (also named cloud ESS or shared ESS) is an ESS that is linked to different end-units and promotes coordination [22]. VESS virtually operates the physical ESS for serving several units. Virtualization reduces the investment cost of ESS and increases its utilization [23,24]. By applying VESS, the limitation of the ESS capacity can be addressed. Moreover, RL is a model-free approach [25]. In recent years, various studies have been conducted using RL for demand response [26], micro-grid systems [27,28], frequency oscillation control [29], and energy trading [30]. The RL-based, model-free approach reduces difficulties related to modeling.

The goal of this study is to design an RL-based VESS operation strategy to manage WPG forecast uncertainty. The VESS logically shares the physical ESS. The suggested VESS operation system considers not only its own operation but also the operation of other units. Under the system, the proposed RL approach to determine VESS operation is presented. Most previous studies used Q-learning-based RL approaches. However, in WPG environments, Q-learning is not efficient owing to the high penetration of uncertainty [31]. In this study, the expected state-action-reward-state-action (SARSA)-based RL approach is applied to obtain a more robust solution to WPG forecast fluctuation. Moreover, multi-dimensional data clustering that considers the WPG forecasting of each unit is combined to enhance the proposed strategy. A simulation study using real datasets recorded by the National Renewable Energy Laboratory (NREL) project of U.S. demonstrates that the proposed strategy provides near-optimal performance.

The main contribution of this study is summarized as follows:

1. This study proposes a VESS operation strategy based on an expected state-action-reward-state-action (SARSA)-based RL approach that is a more robust solution for WPG forecast fluctuation. To the best of our knowledge, this is the first work to apply the RL approach for the VESS operation;
2. The proposed strategy is also combined with multi-dimensional data clustering for enhancing the policy learning performance of the RL approach;
3. Effect of VESS and clustering is carefully discussed, and the usage case of the proposed strategy is suggested.

The rest of this paper is organized as follows. In Section 2, the system description, including forecasting uncertainty model and VESS system, and the VESS operation problem formulation is described, and in Section 3, the design method of the proposed RL-based VESS operation strategy is discussed. In Section 4, measurement studies using real WPG profiles applied to the proposed strategy are demonstrated, and in Section 5, a conclusion of the study is presented.
2. System Description and Problem Formulation

2.1. System Description

2.1.1. Uncertainty Model

In this study, a group of WPGs $\mathcal{U} = \{1, \cdots, u, \cdots, U\}$, such as a wind farm, was considered to be connected to the grid, as shown in Figure 1. To connect to the grid, the WPG operator forecasts the power generation production over $\mathcal{T} = \{1, \cdots, t, \cdots, T\}$, e.g., $T = 24$ h for day-ahead operation. Let $g^u_t$ and $\hat{g}^u_t$ be the actual WPG and its forecasting of the $u$-th WPG at time $t$, respectively. The $u$-th WPG forecast uncertainty at time $t$ is defined as

$$ e^u_t = \hat{g}^u_t - g^u_t. \quad (1) $$

An ESS is operated to manage the uncertainty by charging or discharging energy. By applying the ESS, the uncertainty is calculated as

$$ e^u_t = e^u_t + q^u_t, \quad (2) $$

where $q^u_t$ is the ESS charging/discharging quantity for the $u$ th WPG at time $t$.

Figure 1. Wind power generation system model with a virtual energy storage system; VESS: virtual energy storage system, and RL: reinforcement learning.

2.1.2. VESS System

The ESS is constructed in two parts: a power subsystem (PS) and an energy subsystem (ES). The PS capacity limits instantaneous charging and discharging power. The ES stores the energy, and its capacity determines the ESS service time. The ESS operation is performed within these two constraint regions in the individual ESS [21]. However, the VESS is operated by logically dividing one physical ESS over several units. Therefore, the VESS operation region is limited not only by the ESS capacity of the PS and the ES, but also by the operation of each unit.

Let $a^u_t$ be the ESS charging or discharging action for the $u$ th WPG at decision time $t$. The action is first restricted by the PS capacity $C_{PS}$

$$ -C_{PS} \leq a^u_t \leq C_{PS}, \; \forall \; t \in \mathcal{T}. \quad (3) $$
Moreover, in VESS, the action is limited by the actions of other WPGs. Including this constraint, the action range in Equation (3) is modified as

\[-C_{PS} - \sum_{i \in U / \{u\}} a_i \leq a_u \leq C_{PS} - \sum_{i \in U / \{u\}} a_i, \forall t \in T.\]  

(4)

Considering the charging/discharging efficiency \(\eta\), the actual operation quantity \(q_u\) in Equation (2) is measured as

\[q_u = \begin{cases} \eta a_u, & \text{if } a_u \leq 0, \\ a_u / \eta, & \text{if } a_u > 0. \end{cases}\]  

(5)

Second, the action is performed in the “energy stored” range, referred to as the state-of-charge (SoC). The SoC at the decision time \(t_s\), is measured to be

\[s_t = s_{t-1} + \sum_{u \in U} q_u \Delta T,\]  

(6)

where \(\Delta T\) is the operation time interval, which is limited by the ES capacity \(C_{ES}\),

\[C_{min} \leq s_t \leq C_{max}, \forall t \in T.\]  

(7)

Herein, \(C_{min}\) and \(C_{max}\) express the minimum and maximum operable ES capacity, respectively. These values are determined by considering ESS characteristics, such as the depth of discharge (DoD).

Note that power system requirements such as line flow and frequency should be considered for implementation into practical systems. However, this study focuses on the forecasting error management aspect on the WPG; hence, a simple model used in [23,24] is considered herein. The VESS operation considering power system requirement is an open problem.

2.2. Problem Formulation

The aim of this study is to determine the VESS operation action required to manage WPG forecast uncertainty. Particularly, this study only deals with the WPG forecast error minimization problem. Therefore, the mean absolute error (MAE) is considered to be the performance metric of uncertainty management.

The MAE is calculated as

\[O(a) = \frac{1}{T} \sum_{u \in U} \sum_{t \in T} | e_u + q_u | = \frac{1}{T} \sum_{u \in U} \sum_{t \in T} | e_t |\]  

(8)

where \(a = \{a_1, \cdots, a_u, \cdots, a_T\}\).

Considering the VESS constraints, the VESS operation problem for managing WPG forecast uncertainty can be formulated as

\[
\begin{aligned}
\text{min} \quad & O(a) \\
\text{subject to} \quad & (4) \text{ and } (7).
\end{aligned}
\]  

(9)

If perfect information is known, such as the actual WPG properties about a future time, the problem in Equation (9) can be optimally solved using iteration-based search algorithms such as the gradient descent method and Newton method [32]. However, this non-causal assumption cannot be implemented in the real world. In this study, the solution of this problem using perfect information, including the future time, is used as the optimal solution that is compared to the performance of the proposed VESS operation strategy.
3. Method

As shown in Equation (9), ESS operation is a sequential decision-making (SDM) problem. The SDM problem is mathematically formulated as the state-action space model, and the transaction probability among states is required to optimally solve the problem. However, the RL approach predicts the transaction probability using a learning algorithm, so it requires only the state-action space model to solve the SDM problem in Equation (9) [25].

3.1. State-Action Space

The state-action space for the individually operated ESS is described using a one-dimensional model [31]. The VESS is operated as a physical ESS, so the state-action space for the VESS is also expressed as

\[ S = \{s^1, \ldots, s^l, \ldots, s^\kappa\} \]
\[ A = \{a^0, \ldots, a^\kappa\} \]

where \( a_t \) is the accumulated action for all units \( a_t = \sum_{u\in U} a_t^u \). Considering the PS and ES capacity constraints in Equations (4) and (7), \( \kappa_s \) and \( \kappa_u \) are calculated as

\[ \kappa_s = \left\lfloor \frac{C_{PS} - C_{ES}^{\min}}{C_{ES}} \right\rfloor \]
\[ \kappa_u = \left\lfloor \frac{C_{PS}}{\delta} \right\rfloor \]

where \( \lfloor \cdot \rfloor \) is the floor operation. The discretized ESS operation by the quantization process makes a quantization error, but the error is bound according to the step size [33].

![Figure 2](image-url)  
*Figure 2. State-action space model for a virtual energy storage system; PS: power subsystem, ES: energy subsystem, and ESS: energy storage system.*
At each stage \( t \), the current state is defined as the current SoC, \( s_t \). The next state, \( s_{t+1} \), is determined by the current state and the selected action \( a_{t+1} \) set by Equation (6) as follows:

\[
<s_t = s'_t, a_{t+1} = a'_t> \rightarrow s_{t+1}.
\]

Herein, action \( a_t \) should be selected from within the feasible action range according to the current state \( A_t = \{a_j^\min, \ldots, a_j^1, \ldots, a_j^\max\} \), where \( j^\min = \max(-\kappa_a, 1-i) \) and \( j^\max = \min(\kappa_a, \kappa_s - i) \). As an example, in Figure 2, the feasible action range for the first action \( a_1 \) and the second action \( a_2 \) become \( A_1 = \{a_{-2}, a_{-1}, a_0, a_1, a_2\} \) and \( A_2 = \{a_{-2}, a_{-1}, a_0\} \), respectively.

### 3.2. Decision Policy Design

An RL-based VESS operation strategy determines the decision rule of the current action among the feasible action ranges at each stage in the state-action space. The decision rule is designed to maximize the reward provided by the VESS operation.

The goal of the VESS operation is to minimize the forecast uncertainty presented in the objective function in (9). At stage \( t \), the operator is unaware of the uncertainty of forward time over \( t \). Therefore, the uncertainty included in the VESS operation at stage \( t \) is expressed as

\[
O_t(a_t | \epsilon_t) = \frac{1}{T} \sum_{m=1}^{T} \sum_{u \in U} |\epsilon_u^m|
\]

where the values with a hat represent the expected values.

As shown in Equation (14), the uncertainty comprises the accumulated uncertainty of the current and expected uncertainties at the future time. In the RL approach, the current uncertainty performance of the VESS operation is defined as the instantaneous reward value from the current decision action at each stage. The reward at state \( t \) is presented as

\[
r_t = \frac{1}{T} \sum_{u \in U} |\epsilon_u^t|.
\]

Moreover, the accumulated uncertainty is defined as the return that is the accumulated reward from time \( t \) onward, and is calculated as

\[
R_t = r_t + \gamma r_{t+1} + \cdots + \gamma^{T-t} r_T = r_t + \gamma R_{t+1},
\]

where \( \gamma \) is the discount factor, \((0, 1]\), which reduces the risk of the expected value from the onward decision time. The return in Equation (16) becomes the weighted uncertainty performance of the VESS operation in Equation (14).

The RL-based decision-making approach is used to determine the VESS operation action to minimize the reward, which is the uncertainty performance. For this, the state-action value function is defined and presents the quality of an action, \( a_t \), at a given state \( s_t \), as follows

\[
Q(s_t, a_t) = E[R_t | s_t, a_t] = E[r_t + \gamma Q(s_{t+1}, a_{t+1}) | s_t, a_t].
\]
As the transaction probability of an action at each state is known, \( \pi = \Pr(a_t|s_t) \), \( \forall t \in T, s_t \in S, a_t \in A_t \), the optimal state-action value function \( Q^*(s_t, a_t) \) is measured using the Bellman optimality equation [34]

\[
Q^*(s_t, a_t) = \mathbb{E}[r_t + \gamma \min_{a_{t+1} \in A_{t+1}} Q(s_{t+1}, a_{t+1}) | s_t, a_t]
\]

and the optimal action is determined as

\[
a_t^* = \arg\min_{a_t \in A_t} Q^*(s_t, a_t)
\]

\[
= \arg\min_{a_t \in A_t} \mathbb{E}[r_t|s_t, a_t] + \gamma Q^*(s_{t+1}, a_{t+1}).
\]

However, it is impractical to attempt to determine the transaction probability of an action at each state.

In the RL approach, the state-action value function is estimated by learning. In widely used Q-learning-based RL approaches [26–29], the state-action value function is estimated as

\[
Q^{QL}(s_t, a_t) \leftarrow (1 - \alpha)Q^{QL}(s_t, a_t) + \alpha\left[r_t + \gamma \min_{a_{t+1} \in A_{t+1}} Q^{QL}(s_{t+1}, a_{t+1})\right],
\]

where \( \alpha \) is the learning rate in \((0, 1)\). Moreover, the action is determined as

\[
a_t^{QL} = \arg\min_{a_t \in A_t} \left[r_t + \gamma \min_{a_{t+1} \in A_{t+1}} Q^{QL}(s_{t+1}, a_{t+1})\right].
\]

However, the WPG has a high uncertainty variance [19]. This variance reduces the reliability of the expected value in forward time, such as \( Q^{QL}(s_{t+1}, a_{t+1}) \). Therefore, the Q-learning-based RL approach cannot guarantee uncertainty management performance in WPG environments [31].

Instead of employing the minimum value in the Q-learning-based approach, the expected SARSA-based RL approach uses the expected state-action value to decide the action. The expectation of the value reduces the risk of variance in forward time [25]. In the expected SARSA-based RL approach, the action is determined as

\[
a_t^{SARSA} = \arg\min_{a_t \in A_t} \left[r_t + \gamma \mathbb{E}_{A_{t+1}}\left(Q^{SARSA}(s_{t+1}, a_{t+1})\right)\right].
\]

In addition, the state-action value function is updated as

\[
Q^{SARSA}(s_t, a_t) \leftarrow (1 - \alpha)Q^{SARSA}(s_t, a_t) + \alpha\left[r_t + \gamma \mathbb{E}_{A_{t+1}}\left(Q^{SARSA}(s_{t+1}, a_{t+1})\right)\right].
\]

3.3. Multi-Dimensional Clustering

The determined action for the VESS operation in (22) is an accumulated action set to manage the uncertainty of each WPG, such as \( a_t = [a_t^1, \cdots, a_t^{|A_t|}] \). The expected SARSA approach reduces the expected risk in forward time. However, the multi-dimensional action renders convergence difficult, and also reduces the uncertainty management performance of the VESS operation.

To mitigate this effect of multi-dimensional action, data classification is considered. Data classification is a technique that involves the categorization of data to enable organization for effective operation [35]. With RL approaches, data classification can enhance the learning performance of the state-action value function [36].
In this study, a k-means clustering method is applied, which involves the vector quantization of data into K clusters [37]. With K cluster centroids, \( c = \{c_1, \ldots, c_k, \ldots, c_K\} \), the method classifies the data into K cluster sets, \( C = \{C_1, \ldots, C_k, \ldots, C_K\} \), as follows

\[
\text{argmin}_c \sum_{k=1}^{K} \sum_{u \in U} \| \hat{g}^u - c_k \|_2
\]

where \( \hat{g}^u = \{\hat{g}^u_1, \ldots, \hat{g}^u_T\} \) and \( \| \cdot \|_2 \) express the Euclidean norm. The k-mean clustering problem of (24) is repeatedly solved by the Lloyd algorithm, which determines the centroids of Voronoi diagrams [37].

3.4. RL-Based VESS Operation Strategy

The proposed strategy comprises data clustering for enhancing the learning performance and policy learning to determine the VESS operation action. The proposed strategy is described as follows (Algorithm 1).

**Algorithm 1. Proposed RL-based VESS operation algorithm**

| Step | Description |
|------|-------------|
| 1    | **Initialization** |
| 2    | Set the number of clusters to \( K \). |
| 3    | Initialize centroids \( c_k \) using historical WPG forecasting data. |
| 4    | **Data clustering** |
| 5    | Set cluster \( k \) as \( k = \text{argmin}_k \sum_{u \in U} \| \hat{g}^u - c_k \|_2 \). |
| 6    | Update \( c_k \) including \( \hat{g}^u \). |
| 7    | **Policy learning** |
| 8    | Set \( Q^\text{SARSA}_k \) as \( Q_k = \{Q_1, \ldots, Q_K\} \). |
| 9    | Set \( s_1 \) as the current SoC and \( A_1 \) using (13). |
| 10   | **Policy learning** |
| 11   | **For** \( t = 1, \ldots, T \), |
| 12   | Set \( a^\text{SARSA}_t \) in \( A_t \) using (22). |
| 13   | Update \( s_{t+1}, A_{t+1} \) and \( Q^\text{SARSA} \) using (12) and (23). |
| 14   | **end for** |

First, to apply the k-mean clustering, the number of clusters is set to \( K \), and using the historical WPG forecast data, the centroids are initialized to solve (24) (steps 2 and 3). The cluster number \( k \) of data sets is selected as the cluster that has the minimum Euclidean distance to the cluster centroid in step 5. The cluster number is used to select the active state-action value function for the policy learning process. Moreover, the centroid of the selected cluster is updated considering the dataset in step 6.

Combined with \( K \) clustering, \( K \) state-action value functions are required. In the policy learning process, the k-th state-action value function is loaded as the active state-action value function, \( Q^\text{SARSA}_k \), according to the cluster number \( k \) in step 8. The initial state \( s_1 \) is set as the current VESS condition and the feasible action range \( A_1 \) is determined by the current state \( s_1 \) in step 9. During the operation time horizon \( T \), the VESS operation action is selected using (22) and the values of the state and the state-action value function are updated according to the selected action in steps 12 and 13.

4. Results and Discussion

To verify the performance of the proposed strategy, the simulation results were evaluated. In the simulation, five WPG datasets that were recorded by the National Renewable Energy Laboratory to develop eastern wind resources in the United States of America from 2004 to 2006 were employed [38].
4.1. Performance Results of Proposed Strategy

Figure 3 shows the uncertainty management performance as MAE, with varying VESS size. The black line with circles, the red line with squares, and the blue line with diamonds present the results obtained when applying the optimal solution, the proposed method, and the stochastic method, respectively. The optimal solution is the solution to problem (9), for known information including the future time, and the stochastic method is the VESS operation according to the probabilistic information of the WPG suggested in [40]. The VESS size is the available operation room to manage uncertainty. Therefore, by increasing the size, the MAE is reduced, as shown in Figure 3a. In particular, in Figure 3a the optimal solution and proposed method have a similar slope with increasing VESS size, but the result obtained with the stochastic method shows a less significant decrease. This implies that the optimal solution and the proposed method effectively operate according to the environment, while the stochastic method does not. The stochastic method is designed according to the Markov decision process, similar to the proposed method. However, the stochastic method applying the backward induction approach in [40] predetermines the reserved capacity for the future decision stage according to the probabilistic information of the WPG, so the operational diversity of the stochastic method is lower than those of the optimal solution and proposed method. Figure 3b shows the optimal gap, which represents the difference from the optimal solution. The optimal gap of the proposed method is less than 2%, and is reduced with increasing size. The proposed method can effectively consider environmental characteristics by learning and clustering, and therefore achieves gain by increasing the size. However, the stochastic method cannot reflect this. Therefore, the optimal gap in the stochastic method increases with increasing size.

Each WPG had a capacity of 20 megawatt (MW). Day-ahead forecasting data were provided with 1-h time resolution. Therefore, the operation time horizon was set to 24 h, \( \mathcal{T} = \{1, \cdots, 24\} \).

The simulation results were measured using the data from the first 14 days of December 2006, and the other datasets were used for RL training. Moreover, for policy learning, the learning rate and discount factor were set as \( \alpha = 0.1 \) and \( \gamma = 0.95 \), respectively. The cluster size was assumed to be three. However, a discussion of the cluster size is also presented here.

A lithium-ion based ESS system was considered, which is widely used with renewable energy systems [39]. The charging/discharging efficiency \( \eta \) was set as 0.95, which provided a 0.9 round-trip efficiency, and the DoD margin that restricts the minimum and maximum operable ES capacity was 0.1.

The ES capacity was expressed as the normalized WPG capacity, that is, per unit (p.u.), and the service time was 2 h, which provided a 0.5 charging rate (C-rate).

The simulations in this study were implemented on a 64-bit PC with a 4 GHz Quad-Core Intel Core i7 CPU and 32 GB RAM, using MATLAB R2020a with an IBM CPLEX optimization studio.

4.1. Performance Results of Proposed Strategy

Figure 3 shows the uncertainty management performance as MAE, with varying VESS size. The black line with circles, the red line with squares, and the blue line with diamonds present the results obtained when applying the optimal solution, the proposed method, and the stochastic method, respectively. The optimal solution is the solution to problem (9), for known information including the future time, and the stochastic method is the VESS operation according to the probabilistic information of the WPG suggested in [40]. The VESS size is the available operation room to manage uncertainty. Therefore, by increasing the size, the MAE is reduced, as shown in Figure 3a. In particular, in Figure 3a the optimal solution and proposed method have a similar slope with increasing VESS size, but the result obtained with the stochastic method shows a less significant decrease. This implies that the optimal solution and the proposed method effectively operate according to the environment, while the stochastic method does not. The stochastic method is designed according to the Markov decision process, similar to the proposed method. However, the stochastic method applying the backward induction approach in [40] predetermines the reserved capacity for the future decision stage according to the probabilistic information of the WPG, so the operational diversity of the stochastic method is lower than those of the optimal solution and proposed method. Figure 3b shows the optimal gap, which represents the difference from the optimal solution. The optimal gap of the proposed method is less than 2%, and is reduced with increasing size. The proposed method can effectively consider environmental characteristics by learning and clustering, and therefore achieves gain by increasing the size. However, the stochastic method cannot reflect this. Therefore, the optimal gap in the stochastic method increases with increasing size.

Figure 3. Mean absolute error (MAE) comparison between the optimal solution, proposed method, and stochastic method. (a) Mean absolute error; (b) Optimal gap.
4.2. Effect of VESS

Figure 4 compares the MAE of the individual ESS operation and the proposed VESS operation. The results for individual ESS operations are the optimal solution reformulated problem (9) for each WPG. Each WPG presents different uncertainties, so the decrease in the slope with increasing size also differs, as shown in Figure 4a. However, the results of the proposed VESS operation demonstrate that it outperforms all individual operation results. The individual ESS operation works by using its own information. However, in the case of the proposed VESS operation, information from multiple units is used. Therefore, the proposed VESS operation achieves multi-user diversity gain [41]. Figure 4b verifies the diversity gain. By increasing the size, the operation availability also increases. The proposed VESS operation is effectively operated to achieve availability with multi-user diversity. Therefore, the VESS operation gain compared to the individual ESS operation is enhanced with increasing operation availability.

![Figure 4.](image)

(a) (b)

Figure 4. MAE comparison among individual ESS operation and VESS operation. (a) Mean absolute error; (b) Optimal gap.

4.3. Effect of Clustering

Figure 5 shows the optimal gap of the proposed method with 1, 3, and 5-cluster cases. As shown in Figure 5, the optimal gap is reduced as the number of clusters increases. In particular, with five clusters, an optimal gap enhancement of more than 1.5% can be obtained when the ESS size is 0.6 p.u. This indicates that the clustering method is an effective way to enhance the performance of the proposed method. However, compared to the performance enhancement provided by three clusters or one cluster, the performance increase provided by five clusters is less than that of three clusters. This is because the distance between the centroids is reduced with increasing cluster size. Moreover, by increasing the cluster size, the number of state-action value functions also increases for policy learning. This increases the system complexity in the implementation. Therefore, it is important to set the appropriate cluster size by considering both performance enhancement and system complexity. As an example, in this study operating with five WPG units, three clusters are efficient considering the performance enhancement, as shown in Figure 5.
WPG sets as 1.5, in case of the individual ESS operation, each ESS size larger than 1 p.u. is required, VESS service with economic benefit by reducing the ESS size. Moreover, by increasing the number of clusters, the ESS size can be reduced, as shown in Figure 5. This is economically not viable. However, in the proposed VESS case, 0.2 p.u. of ESS size is required for each WPG with the same target. This makes a business model such as a VESS service with economic benefit by reducing the ESS size. Moreover, by increasing the number of clusters, the ESS size can be reduced, as shown in Figure 5. The cluster size affects the number of the state-action value function that is related to the memory size and the computational complex. Therefore, the VESS service provider can select the ESS size and the number of clusters considering the ESS cost, the memory cost, and the computational complexity, as well as the WPG forecast error management target.

4.4. Usage of the Proposed Strategy

The VESS operation applying the proposed strategy can get a higher forecast error management performance than that of the individual ESS operation. For an example, when the MAE target of each WPG sets as 1.5, in case of the individual ESS operation, each ESS size larger than 1 p.u. is required, as shown in Figure 4a. This is economically not viable. However, in the proposed VESS case, 0.2 p.u. of ESS size is required for each WPG with the same target. This makes a business model such as a VESS service with economic benefit by reducing the ESS size. Moreover, by increasing the number of clusters, the ESS size can be reduced, as shown in Figure 5. The cluster size affects the number of the state-action value function that is related to the memory size and the computational complex. Therefore, the VESS service provider can select the ESS size and the number of clusters considering the ESS cost, the memory cost, and the computational complexity, as well as the WPG forecast error management target.

5. Conclusions

This study proposed an RL-based VESS operation strategy to manage WPG forecasting uncertainty. The VESS operation model is the first to consider not only its own uncertainty management requirement, but also the requirements of other units. Applying the VESS model, the expected SARSA-based learning policy is suggested to solve the sequential decision-making problem of the VESS operation. Moreover, the k-means data clustering method is employed to enhance the performance of the proposed strategy by reducing uncertainty variance. The simulation results demonstrate that the proposed strategy provides a near-optimal performance, with a less than 2%-point gap to the optimal solution that requires information including the future time. Moreover, the MAE improvement when applying the proposed method has a similar slope to that of the optimal method according to the storage size. This shows that the proposed method obtains a similar operational diversity to that of the optimal method and can achieve near-optimal performance generally. In addition, we evaluated the performance achieved by the VESS operation in terms of multi-user diversity and the effect of the clustering method according to cluster size.

Research on VESS operation is at an early stage. This study shows that VESS operation can outperform individual ESS operation. However, the performance enhancement according to the VESS operation differs for each unit. Therefore, the VESS operation considering the performance balance among units will be the subject of further research. Moreover, this study only considers a simple system model. By including power system requirements, the system model can be practically extended further. Finally, the forecast error management of WPG is highly related to the revenue, and the VESS operation.
Applied Sciences, 10, 6420

is cost-efficient, rather than the individual ESS operation. Therefore, this study can be extended to research in the economic aspect, such as a revenue maximization problem considering ESS costs.

**Funding:** This work was supported by 2020 Research Grant of Hanseo University.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. International Energy Agency (IEA). World Energy Balances. Statistics Report—July 2020. Available online: https://www.iea.org/reports/world-energy-balances-overview (accessed on 31 August 2020).
2. REN21. Renewables 2020 Global Status Report. June 2020. Available online: https://www.ren21.net/gsr-2020 (accessed on 31 August 2020).
3. International Energy Agency (IEA). Renewables 2019. Fuel Report—October 2019. Available online: https://www.iea.org/reports/renewables-2019 (accessed on 31 August 2020).
4. Ayodele, T.R.; Jimoh, A.; Munda, J.L.; Tehile, A.J. Challenges of grid integration of wind power on power system grid integrity: A review. Int. J. Renew. Energy Res. 2012, 2, 618–626.
5. Wu, Y.-K.; Chang, S.-M.; Mandal, P. Grid-connected wind power plants: A survey on the integration requirements in modern grid codes. IEEE Trans. Ind. Appl. 2019, 55, 5584–5593. [CrossRef]
6. Hao, Y.; Tian, C. A novel two-stage forecasting model based on error factor and ensemble method for multi-step wind power forecasting. Appl. Energy 2019, 238, 368–383. [CrossRef]
7. Sun, M.; Feng, C.; Zhang, J. Conditional aggregated probabilistic wind power forecasting based on spatio-temporal correlation. Appl. Energy 2019, 256, 113842. [CrossRef]
8. Demolli, H.; Dokuz, A.; Şakir Ecemis, A.; Gökçek, M. Wind power forecasting based on daily wind speed data using machine learning algorithms. Energy Convers. Manag. 2019, 198, 111823. [CrossRef]
9. Zhao, H.; Wu, Q.; Hu, S.; Xu, H.; Rasmussen, C.N. Review of energy storage system for wind power integration support. Appl. Energy 2015, 137, 545–553. [CrossRef]
10. Sebastian, R. Application of a battery energy storage for frequency regulation and peak shaving in a wind diesel power system. IET Gener. Transm. Distrib. 2016, 10, 764–770. [CrossRef]
11. Atherton, J.; Sharma, R.; Salgado, J. Techno-economic analysis of energy storage systems for application in wind farms. Energy 2017, 135, 540–552. [CrossRef]
12. Bitaraf, H.; Rahman, S. Reducing curtailed wind energy through energy storage and demand response. IEEE Trans. Sustain. Energy 2018, 9, 228–236. [CrossRef]
13. Gomes, I.; Pousinho, H.M.I.; Melício, R.; Mendes, V. Stochastic coordination of joint wind and photovoltaic systems with energy storage in day-ahead market. Energy 2017, 124, 310–320. [CrossRef]
14. Sperstad, I.B.; Korpås, M. Energy storage scheduling in distribution systems considering wind and photovoltaic generation uncertainties. Energies 2019, 12, 1231. [CrossRef]
15. Kalavani, F.; Mohammadi-Ivatloo, B.; Zare, K. Optimal stochastic scheduling of cryogenic energy storage with wind power in the presence of a demand response program. Renew. Energy 2019, 130, 268–280. [CrossRef]
16. Yuan, Y.; Zhang, X.; Ju, P.; Li, Q.; Qian, K.; Fu, Z. Determination of economic dispatch of wind farm-battery energy storage system using Genetic Algorithm. Int. Trans. Electr. Energy Syst. 2012, 24, 264–280. [CrossRef]
17. Khare, V.; Nema, S.; Baredar, P. Optimisation of the hybrid renewable energy system by HOMER, PSO and CPSO for the study area. Int. J. Sustain. Energy 2015, 36, 1–18. [CrossRef]
18. Liu, Z.-F.; Li, L.; Tseng, M.-L.; Tan, R.; Aviso, K. Improving the reliability of photovoltaic and wind power storage systems using least squares support vector machine optimized by improved chicken swarm algorithm. Appl. Sci. 2019, 9, 3788. [CrossRef]
19. Oh, E.; Son, S.-Y. Energy-storage system sizing and operation strategies based on discrete Fourier transform for reliable wind-power generation. Renew. Energy 2018, 116, 786–794. [CrossRef]
20. Dui, X.; Zhu, G.; Yao, L. Two-stage optimization of battery energy storage capacity to decrease wind power curtailment in grid-connected wind farms. IEEE Trans. Power Syst. 2018, 33, 3296–3305. [CrossRef]
21. Oh, E.; Son, S.-Y. Theoretical energy storage system sizing method and performance analysis for wind power forecast uncertainty management. Renew. Energy 2020, 155, 1060–1069. [CrossRef]
22. Li, S.; Yang, J.; Fang, J.; Liu, Z.; Zhang, H. Electricity scheduling optimisation based on energy cloud for residential microgrids. IET Renew. Power Gener. 2019, 13, 1105–1114. [CrossRef]
23. Oh, E.; Son, S.-Y. Shared electrical energy storage service model and strategy for apartment-type factory buildings. *IEEE Access* 2019, 7, 130340–130351. [CrossRef]
24. Zhao, D.; Wang, H.; Huang, J.; Lin, X. Virtual energy storage sharing and capacity allocation. *IEEE Trans. Smart Grid* 2020, 11, 1112–1123. [CrossRef]
25. Sutton, R.; Barto, A. Reinforcement learning: An introduction. *IEEE Trans. Neural Netw.* 1998, 9, 1054. [CrossRef]
26. Vázquez-Canteli, J.R.; Nagy, Z. Reinforcement learning for demand response: A review of algorithms and modeling techniques. *Appl. Energy* 2019, 235, 1072–1089. [CrossRef]
27. Qiu, X.; Nguyen, T.A.; Crow, M.L. Heterogeneous energy storage optimization for microgrids. *IEEE Trans. Smart Grid* 2015, 7, 1453–1461. [CrossRef]
28. Bui, V.-H.; Hussain, A.; Bui, V.-H. Q-Learning-based operation strategy for Community Battery Energy Storage System (CBESS) in microgrid system. *Energies* 2019, 12, 1789. [CrossRef]
29. Younesi, A.; Shayeghi, H. Q-learning based supervisory pid controller for damping frequency oscillations in a hybrid mini/micro-grid. *Iran. J. Electr. Electron. Eng.* 2019, 15, 126–141.
30. Wang, H.; Huang, T.; Liao, X.; Abu-Rub, H.; Chen, G. Reinforcement learning in energy trading game among smart microgrids. *IEEE Trans. Ind. Electron.* 2016, 63, 1. [CrossRef]
31. Oh, E.; Wang, H. Reinforcement-learning-based energy storage system operation strategies to manage wind power forecast uncertainty. *IEEE Access* 2020, 8, 20965–20976. [CrossRef]
32. Boyd, S.P.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.
33. Gray, R.M.; Neuhoff, D.L. Quantization. *IEEE Trans. Inf. Theory* 1998, 44, 2325–2383. [CrossRef]
34. Kirk, D.E. *Optimal Control Theory: An Introduction*; Prentice Hall: Upper Saddle River, NJ, USA, 1970.
35. Durgesh, K.S.; Lekha, B. Data classification using support vector machine. *J. Theor. Appl. Inf. Technol.* 2010, 12, 1–7.
36. Rajan, S.; Ghosh, J.; Crawford, M. An active learning approach to hyperspectral data classification. *IEEE Trans. Geosci. Remote. Sens.* 2008, 46, 1231–1242. [CrossRef]
37. Lloyd, S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* 1982, 28, 129–137. [CrossRef]
38. Brower, M. Development of Eastern Regional Wind Resource and Wind Plant Output Datasets: 3 March 2008–31 March 2010; Office of Scientific and Technical Information (OSTI): Washington, DC, USA, 2009.
39. Yang, Y.; Bremner, S.; Menictas, C.; Kay, M. Battery energy storage system size determination in renewable energy systems: A review. *Renew. Sustain. Energy Rev.* 2018, 91, 109–125. [CrossRef]
40. Oh, E.; Son, S.-Y.; Hwang, H.M.; Park, J.-B.; Lee, K.Y. Impact of demand and price uncertainties on customer-side energy storage system operation with peak load limitation. *Electr. Power Compon. Syst.* 2015, 43, 1872–1881. [CrossRef]
41. Verdú, S. *Multiuser Detection*; Cambridge University Press: Cambridge, UK, 1998.