Optimizing Bidding Strategies of a Combined Wind/ESS System Using a Novel Stochastic Optimization with Reinforcement Learning Algorithm

Anne Stratman, Wei Qiao, Liyan Qu

University of Nebraska – Lincoln, Dept. of Electrical Engineering

Abstract

Abstract: As the amount of wind power in the global energy supply grows, increasing reliability between the cleared day-ahead (DA) bids and the real-time (RT) output of wind power producers (WPPs) will be essential for ensuring power system stability. However, DA wind power forecasts are highly uncertain, and the RT wind production is often less than the DA forecast, causing the WPP to default on its cleared DA offer. Energy storage systems (ESSs) can be used to improve wind power reliability by reducing shortages from the WPP's cleared DA offer. The ESS can strategically charge when the cost of electricity is low, and discharge when there is a shortage between the WPP's cleared DA offer and the RT wind production. This work proposes a novel method of finding the optimal ESS strategy and wind power bids. A traditional technique, stochastic optimization is used to find the offer strategy of the WPP. Q-learning, which is a type of reinforcement learning, is used to find the optimal ESS strategy. The stochastic optimization with reinforcement learning (SORL) approach is compared to a traditional method where stochastic optimization is used to find the strategy of both the wind bidding and the ESS. Case study results show that the SORL method reduces RT deviations from the WPP's cleared DA amount compared to the traditional stochastic optimization method.

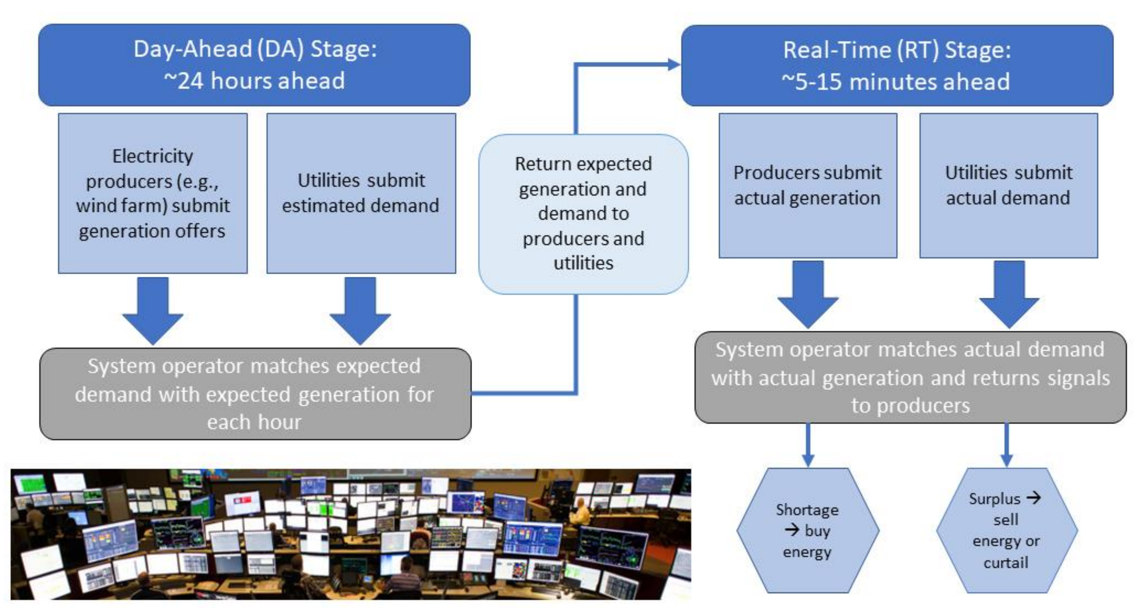

Fig. 1: Electricity market operations for the Southwest Power Pool (SPP)

Image source: https://www.spp.org/markets-operations/market-monitoring/

Fig. 1: Bidding stages in US electricity markets.

Case Study

The proposed SORL algorithm is tested on a hybrid wind-ESS system consisting of an 80-MW wind farm and a 20-MW ESS. Data for DA and RT LMPs are obtained from the PJM Electricity Market [1]. Wind power data is obtained from NREL's Wind Toolkit [2]. Four years of data are used to train and test the model – the first three years and eight months of data are used to train the SORL algorithm, while the final four months of data are used for testing. We assume a minimum offer requirement of 5MW and that the WPP's DA offer cannot be greater than the wind power forecast, based on PJM policy. We also assume that the WPP's entire DA or RT offer is always accepted.

Optimization Setup

Initialization:

Generate wind power forecasts by adding random error to historical data. Separate forecasts into clusters by forecast amount.

Define ESS actions: charge or discharge by 0, 1, 2, 3, 4, or 5 MW, if the action does not cause an SOC < 20% / > 100%

Create a Q-learning table for each wind power cluster.

Optimization Process:

Choose ESS action.

Use stochastic optimization to find the DA wind power offer with equation (1).

Calculate profit of each hour by substituting the true historical wind power value for P_F.

Calculate Q-learning reward for each hour by comparing the true profit from the SORL model to the true profit from the traditional stochastic optimization model.

Update the Q-learning table using (3).

$$\max\{\lambda_{DA}(P_{DA} - a_{C,DA}) + \lambda_{RT}(\Delta^{+} - \Delta^{-})\} + \beta\left(\zeta - \left(\frac{1}{1-A}\right)\varphi\right) \quad (1)$$

$$\Delta^{+} - \Delta^{-} = P_F - P_{DA} - a_{C,RT} - \frac{1}{\eta}a_{D,RT} \quad (2)$$

$$Q^{new}(s_t, a_t) = (1-\alpha)\cdot Q(s_t, a_t) + \alpha\left(r_t + \gamma\cdot\max_{a} Q(s_{t+1}, a)\right) \quad (3)$$

Results

As shown in Table I, the total offer amounts for the two methods are nearly identical, with the difference in total offer values caused by roundoff error. Introducing the ESS allows the WPP to compensate for most forecast errors by discharging during hours with a RT shortage, leading to relatively low shortage amounts compared to the total amount of DA offers. When the SORL method is used, the WPP has 1447 MW less in RT shortages compared to the stochastic optimization method, which is a 52.7% reduction.

Table I: Shortage from the WPP's cleared DA offer, for the SORL and stochastic optimization methods.

Method	Total Offer (MW)	Shortage (MW)
Stochastic opt.	63,924	2747
SORL	63,928	1300

Under the stochastic optimization method, the ESS tends to charge for several hours in a row when prices are low, then discharge while prices are high (Fig. 2). Under the SORL method, the ESS alternates between charging and discharging much more frequently. The stochastic optimization method uses the entire SOC range, while with the SORL method the maximum SOC is ~70%.

Since discharging the ESS allows the WPP to sell power in RT or compensate for a shortage from the DA wind bid, increasing the WPP's profit, the optimal action for each state involves discharging the ESS, as shown in Table II.

Table II: Optimal ESS actions.

SOC	Low Wind Power	High Wind Power
20-40%	Discharge 2MW	Discharge 1MW
40-60%	Discharge 5MW	Discharge 3MW
60-80%	Discharge 3MW	Discharge 5MW
80-100%	Discharge 3MW	Discharge 5MW

In Table III, the expected DA profit is the expected profit based on the wind power forecast and DA bid. The RT profit is the profit in the RT market from making a RT wind offer or selling surplus wind or ESS power. RT charges are penalties that the WPP must pay in RT for shortages from its DA bid, or the cost of charging the ESS. Finally, the net profit is the sum of the DA and RT profits given the ESS action, DA wind bid, and the RT wind production. The SORL method reduces RT charges by 41.76% compared to the stochastic optimization method.

Table III: Expected DA profit, RT profit, RT charges, and net profit over the testing period.

Method	DA Expected Profit ($)	RT Profit ($)	RT Charges ($)	Net Profit ($)
Stochastic opt.	2,847,000	2,595,900	101,000	5,342,000
SORL	2,817,100	2,426,600	58,814	5,184,900

Conclusions

The proposed SORL algorithm successfully reduces the RT shortage from the WPP's cleared DA bid by 52.7% compared to the stochastic optimization method. While stochastic optimization finds the best ESS action for the current timestep, the SORL algorithm uses Q-learning to consider the impact of the present hour's ESS action on future profit. This allows the ESS to learn a strategy that will maintain a sufficient SOC to help the WPP avoid penalties from default in the future. With the stochastic optimization method, since the ESS frequently discharges to the minimum SOC, it cannot be used to avoid default penalties during those hours, and the WPP faces higher RT charges than with the SORL method. The proposed SORL method is beneficial to both WPPs, who will face lower penalties for RT shortages from their cleared DA bids; and to system operators, who will benefit from greater consistency between the DA forecasted and RT wind power output.

Future work will involve using deep reinforcement learning the model with continuous ESS actions to overcome this limitation. Future work will also involve including the DA and RT LMPs and wind power forecast in the Q-learning states to find the optimal ESS strategy considering the electricity prices.

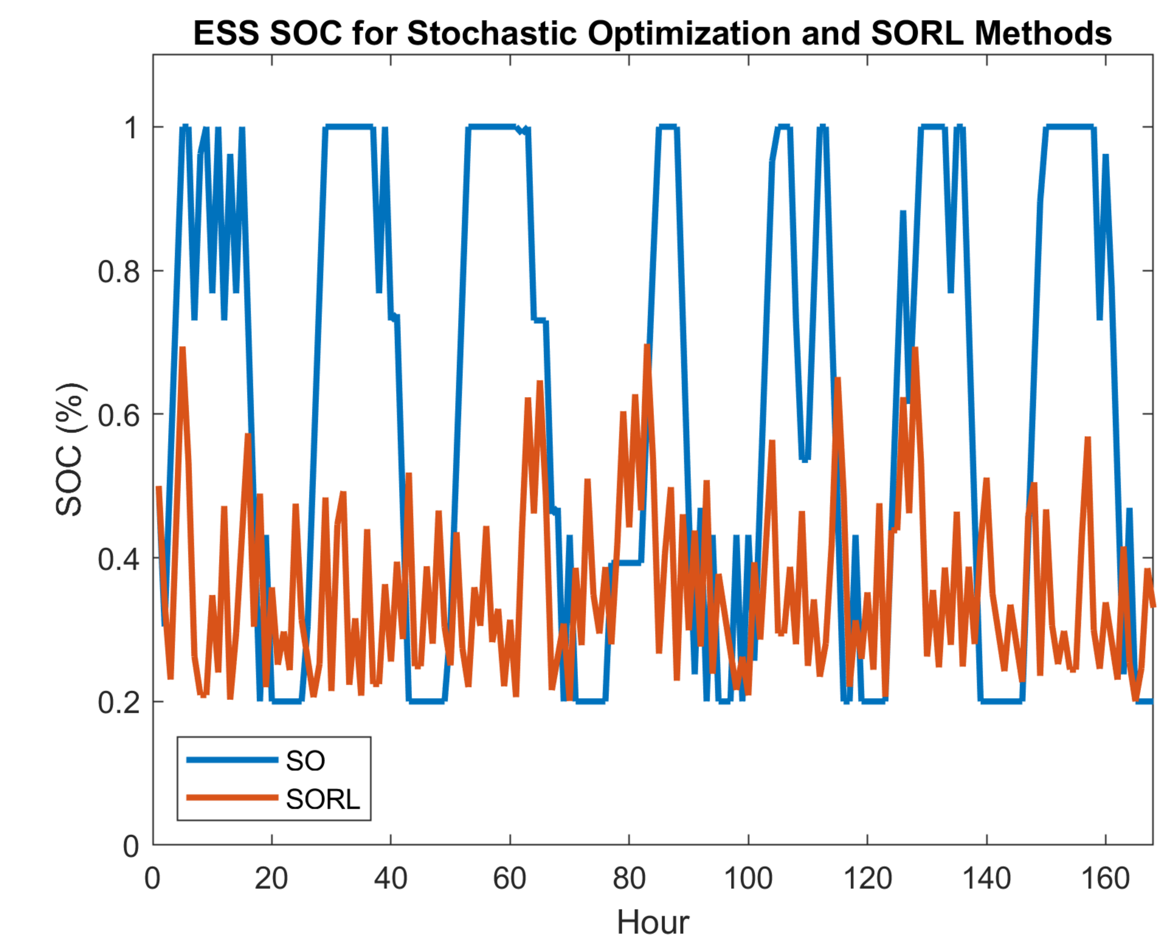

Fig. 2: ESS SOC during each hour of one week of the testing period for the stochastic optimization (blue) and SORL (red) methods.

Funding Acknowledgment

This work was supported by the Nebraska Public Power District through the Nebraska Center for Energy Sciences Research at the University of Nebraska-Lincoln.