Reinforcement Learning for Optimal a Hour-Ahead Electricity Trading with Battery Storage

Peyman Kor[1]

[1] Department of Energy Resources, University of Stavanger, Stavanger, Norway

Introduction

- The coastal region of Europe has seen a growing interest in wind energy. A challenge with wind energy sources is handling the variability.
- A simple solution to handle this variability is to sell any electricity from a renewable source into the grid and use the tremendous power of the grid to handle this variability.
- However, there has been considerable interest in using storage (particularly battery storage) to handle variability in renewable energies. The battery storage provides flexibility to the energy provider.

Objectives

1. Illustrate and introduce the workflow for solving sequential Decision problems
2. Present Value Function Approximation (VFA) based on the Bellman optimality equation as a policy for decision in hand.

Figure 1: The Electricity Storage Decision Schematics

Methods

`Lookahead Policy` is optimal:

$$X_t^*(S_t) = \arg\max_{x_t}\left(C(S_t,x_t) + \mathrm{E}\left\{\max_{\pi\in\Pi}\left\{\mathrm{E}\sum_{t'=t+1}^{T} C(S_{t'},X_{t'}^{\pi}(S_{t'}))\mid S_{t+1}\right\}\mid S_t,x_t\right\}\right)$$

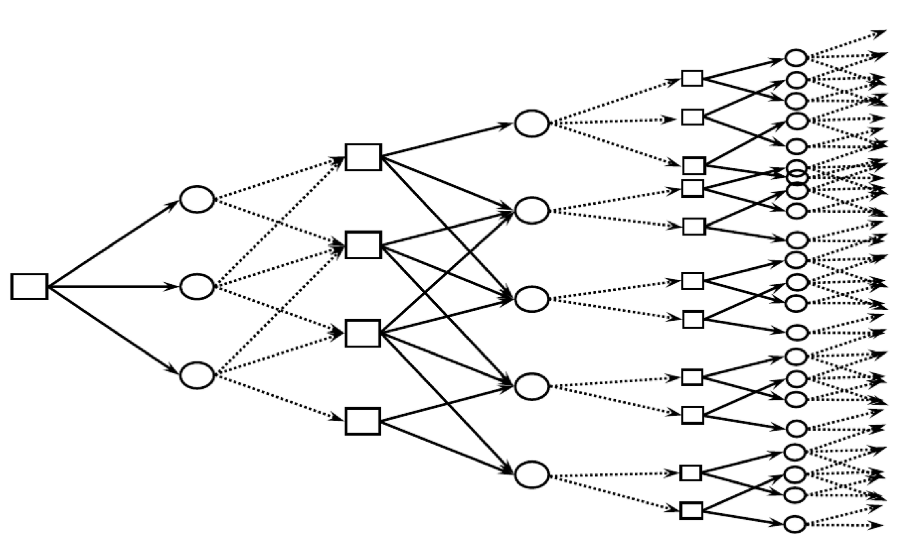
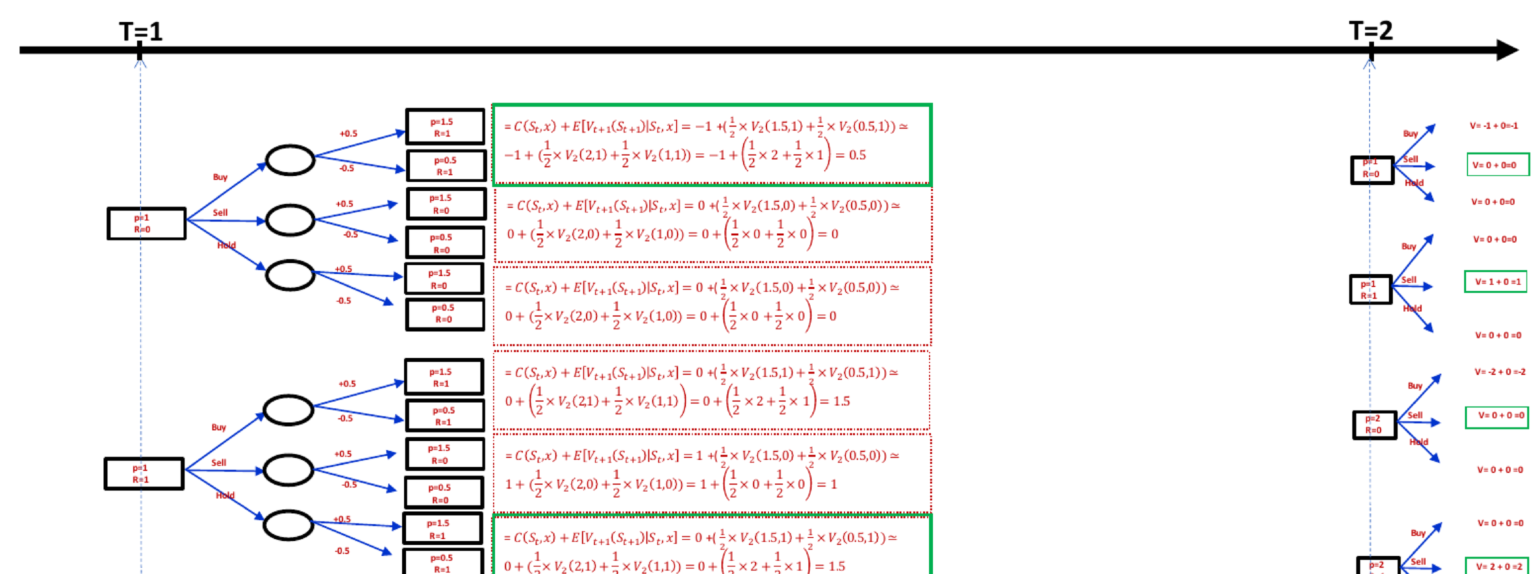


Figure 2: Lookahead Policy as Decision Tree

Results

Visual Illustration of Concept:

- Battery could be fully charged Or Discharged, R t = [1,0]
- The Electricity Price could have three possible scenario, p=[1,2,3]
- The Electricity price could decrease or increase with ,Δp = [−0.5,0.5] each with chance.
- The number of time steps is 3, T = [0,1,2]



T = 1

State (#)	P (price)	R (Charged, Dischareged)	V(p,R)
1	1	0	0.5
2	2	0	0.5
3	3	0	0
4	1	1	1.5
5	2	1	2.5
6	3	1	3

T = 2

State (#)	P (price)	R (Charged, Dischareged)	V(p,R)
1	1	0	0
2	2	0	0
3	3	0	0
4	1	1	1
5	2	1	2
6	3	1	3

Figure 3: Lookahead Policy Computation, Small Example

Two data sets, price on Week 1 2005 as train, price on Week 1 2006 to test:

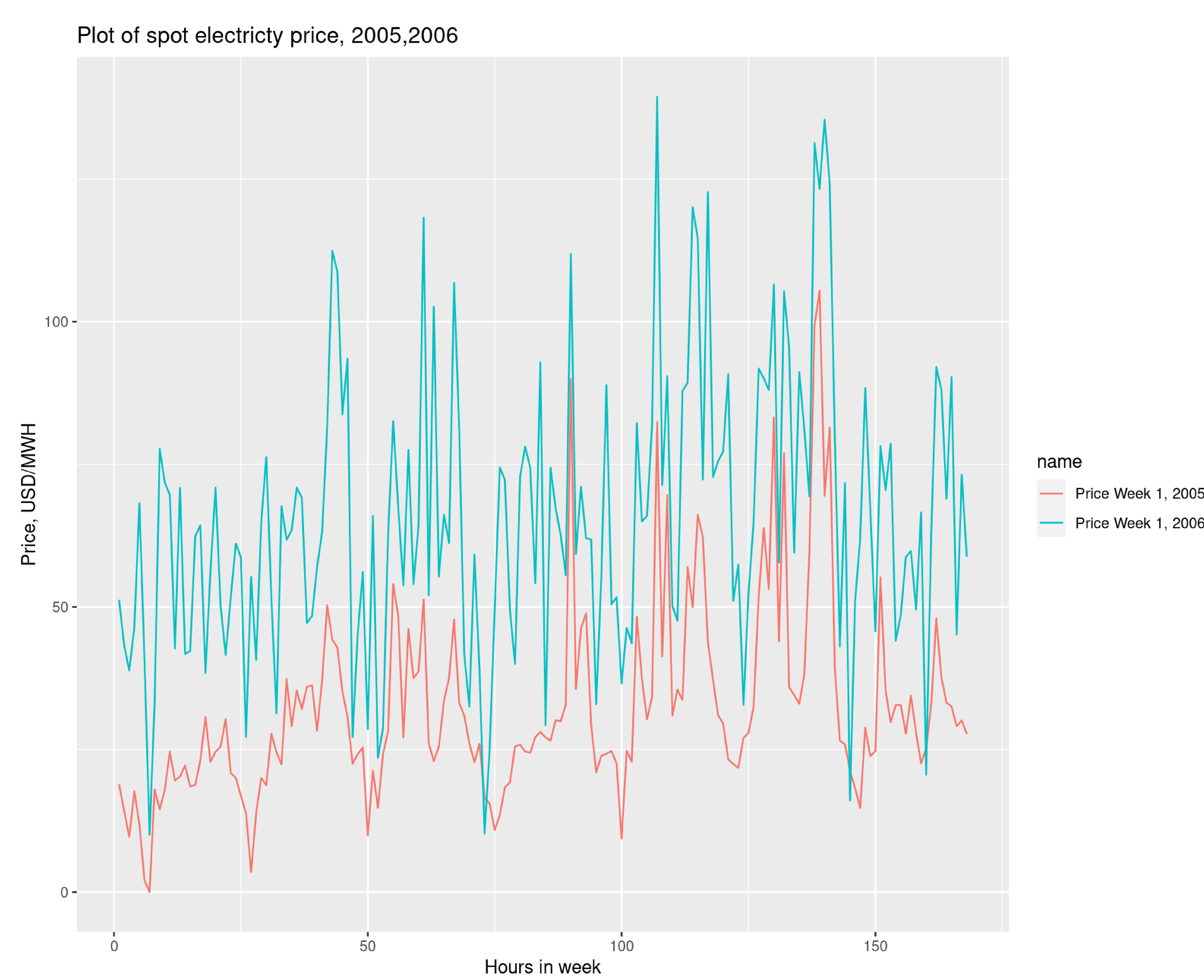


Figure 4: Electricity price at same week, two years

Distribution of Hourly Price Change in Training data:

- The change in price is needed to model uncertainty in price.

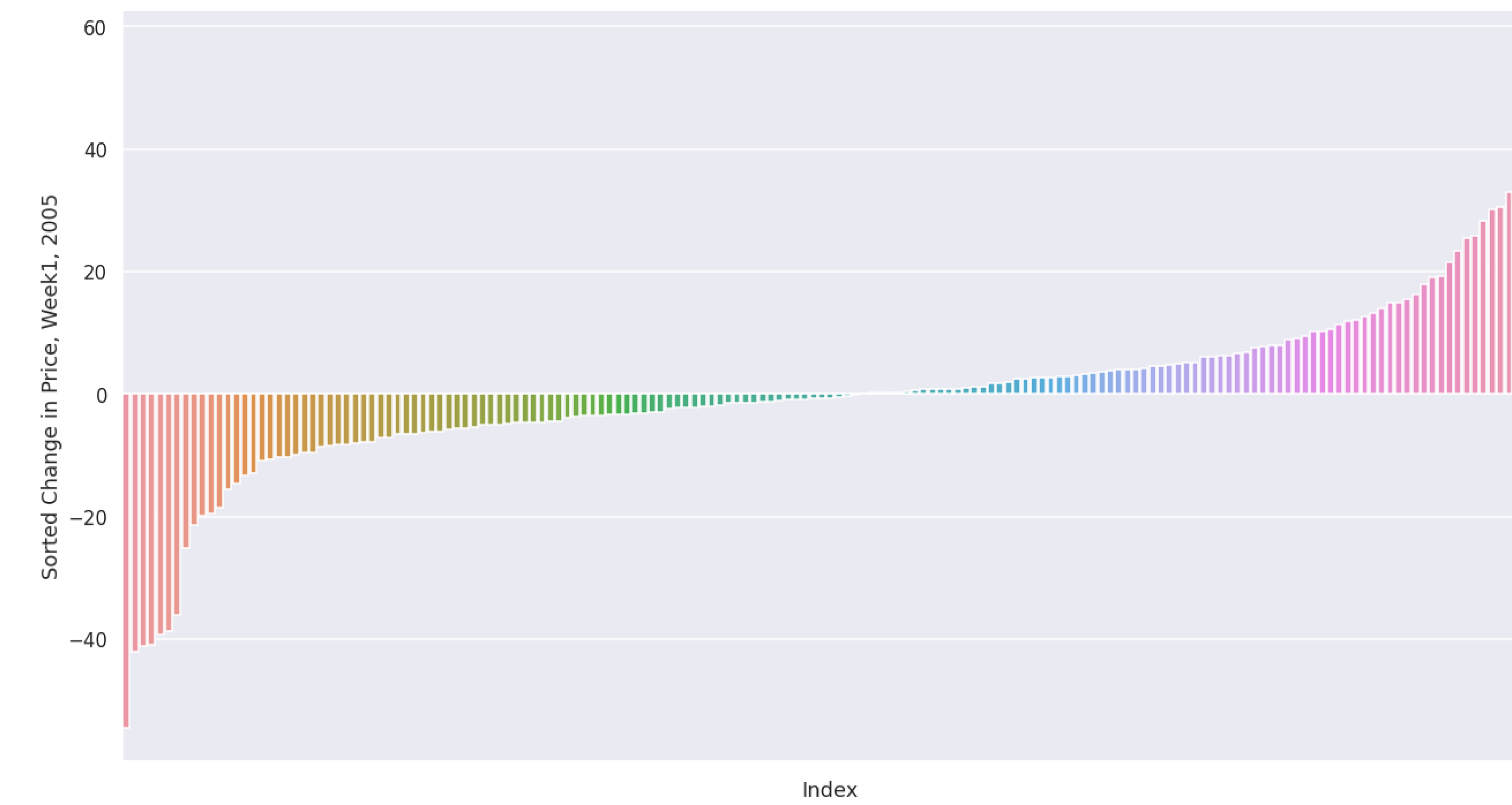


Figure 5: Sort of change in hourly price

How stored electricity is traded (sell and buy to grid).

- At each time step, the decision is made with taking into account all the `Future` decisions:

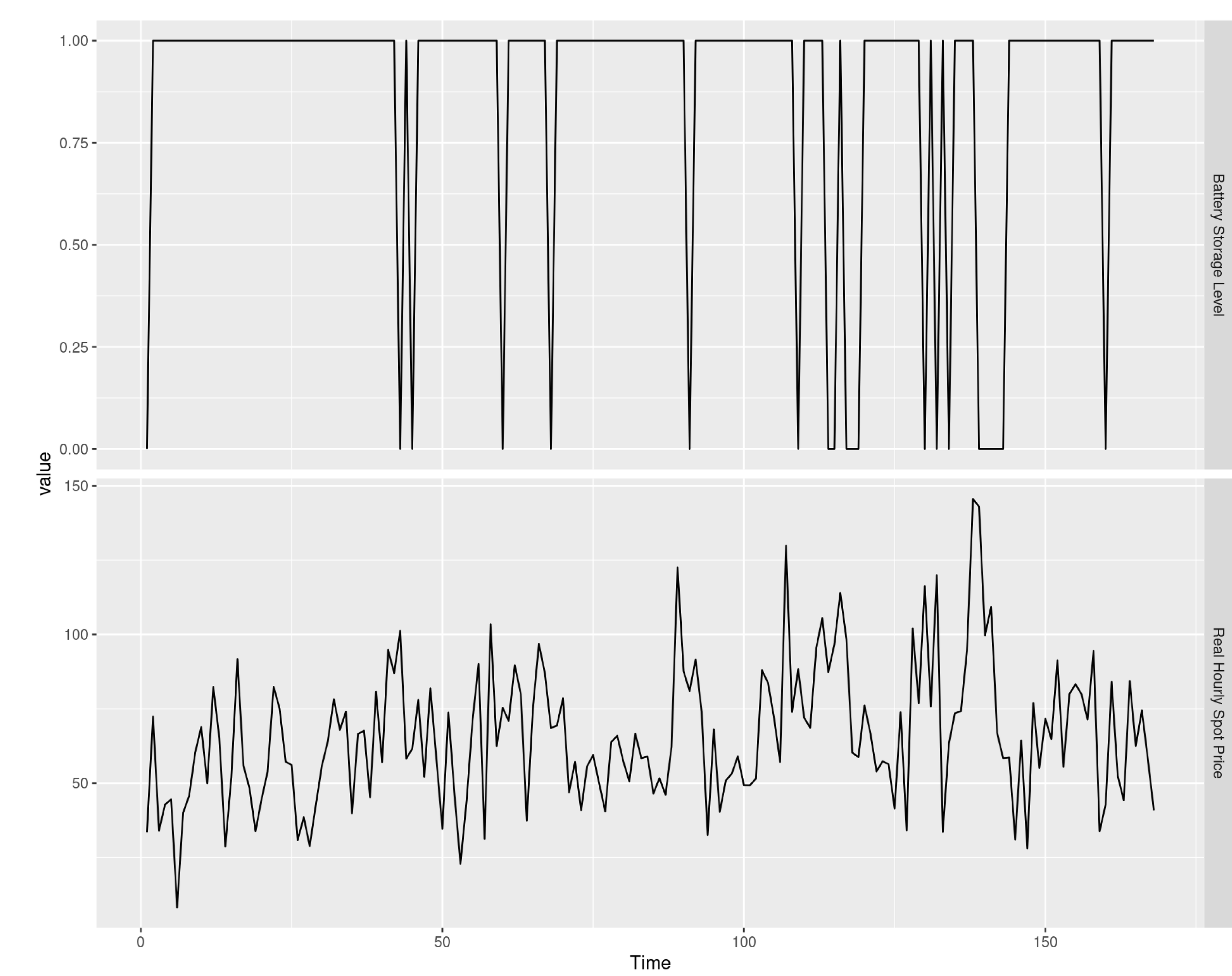


Figure 6: Buying and Selling Time of Battery and, Electricity Price in lower

- The final contribution (profit) with 1 MWh, is `516.3` USD per week.

Conclusion

- We show that policy trained using Lookahead policy on historical price data consistently generated profit revenue for decision-maker.
- The next step is to apply workflow for more difficult, stochastic spot electricity prices.