

## 一、模型假设

不考虑需求响应

## 二、模型建立

### 2.1 优化目标

确定合适的策略  $\Pi$ ，也就是每个 DBESS 在时间  $1-T$  上的充放电功率序列  $P$  使得分布式储能的聚合商的  $n$  个 DBESS 在  $T$  个周期内利益最大。

$$\max_{\pi} w_{comm} \sum_{i=1}^n \sum_{t=1}^T P_{i,t} \Delta t \rho_t - \sum_{t=1}^T C_t$$

在不考虑需求响应的情况下，由于每个 DBESS 均为独立运行互不干扰，所以优化目标可以简化为单个 DBESS 的收益最大。

### 2.2 约束条件

#### 2.2.1 DBESS 充放电约束

充放电功率小于额定功率

$$|P_i| \leq P_{iN}$$

有足够容量放电，充电不能超过额定电量

$$\begin{cases} P_i \Delta t < S_{i,t-1} \\ -P_i \Delta t + S_{i,t-1} < Q_{Ni} \end{cases}$$

状态转换方程

$$S_{i,t} = S_{i,t-1} + P_i \Delta t$$

#### 2.2.2 调用成本

充电损耗成本

$$C_{i,char} = \rho_t P_i (1 - \eta_i) \Delta t$$

放电损耗成本

$$C_{i,dch} = \rho_t P_i \left( \frac{1}{\eta_i} - 1 \right) \Delta t$$

老化成本

$$C_{i,ag} = \frac{(d_{k,i} + \frac{P_i}{E_i})^{K_p} - d_{k,i}^{K_p}}{2N} C_{i,ap}$$

### 三、模型求解

由于同类 DBESS（容量与额定功率相同）收敛结果一定相同，因此在算例中  $i$  代表不同类的 DBESS，同类 DBESS 均只考虑其中一个储能电源的充放电策略。

#### 3.1 强化学习参数定义

参数	含义	定义式
$\Delta P$	最小的充放电功率（步长）	$\Delta P = 2$
$S_{i,t}$	第 $i$ 类 DBESS 在 $t$ 时刻的荷电状态	$S_{i,t} = [0, \Delta P \Delta t, 2\Delta P \Delta t, \dots, Q_N]$
$a_{i,t}$	第 $i$ 类 DBESS 在 $t$ 时刻的动作	$a_{i,t} = [-Q_N, -Q_N + \Delta P, \dots, 0, \Delta P, \dots, Q_N]$
$R_{i,t}$	第 $i$ 类 DBESS 在 $t$ 时刻的奖励	$R_{i,t} = P_i \Delta t \rho_t$
$Q_{i,t}$	第 $i$ 类 DBESS 在 $t$ 时刻的值函数矩阵	$Q_{i,t} = \{S_{i,t}, a_{i,t}\}$

#### 3.2 强化学习方法

Q-Learning 算法，行动策略采用 eps-greedy 策略，目标策略为贪心策略。值函数更新方法为

$$Q\{i,t\}(S_{i,t},a_{i,t}) \leftarrow Q\{i,t\}(S_{i,t},a_{i,t}) + \alpha(\lambda \max_{a_{i,t+1}} Q\{i,t+1\}(S_{i,t+1},a_{i,t+1}) - Q\{i,t\}(S_{i,t},a_{i,t})) - C_{i,char/dch}$$

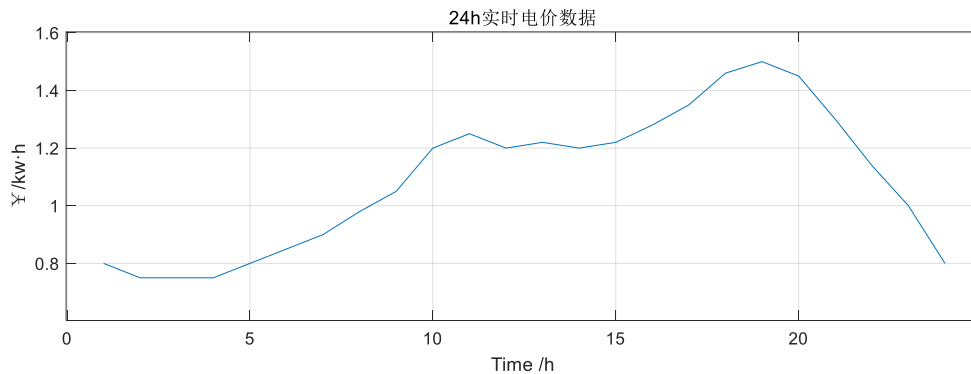
### 四、算例分析

#### 4.1 算例一 单个 DBESS 充放电策略

不考虑老化成本，且只考虑一类 DBESS，观察一天内 DBESS 的充放电策略与电价的关系。

##### 4.1.1 算例参数

##### 1. 电价参数



##### 2. DBESS 数据

电池类型	磷酸铁锂
额定容量 $Q_N$	60 kW · h
额定充放电功率 P	30 kW
初始荷电状态 SOC	50%

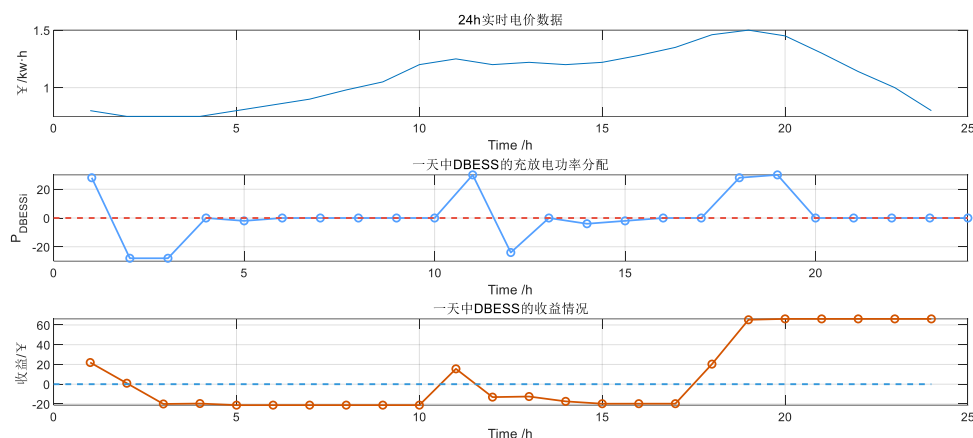
初始效率  $\eta$

98%

运行周期 T

24 h

#### 4.1.2 算例结果



DBESS 在经过足够次数的学习之后，充放电策略趋于低储高发，单个 DBESS 在 24h 累积收益超过 60 元。

#### 4.2 算例二 不同老化程度 DBESS 收益对比

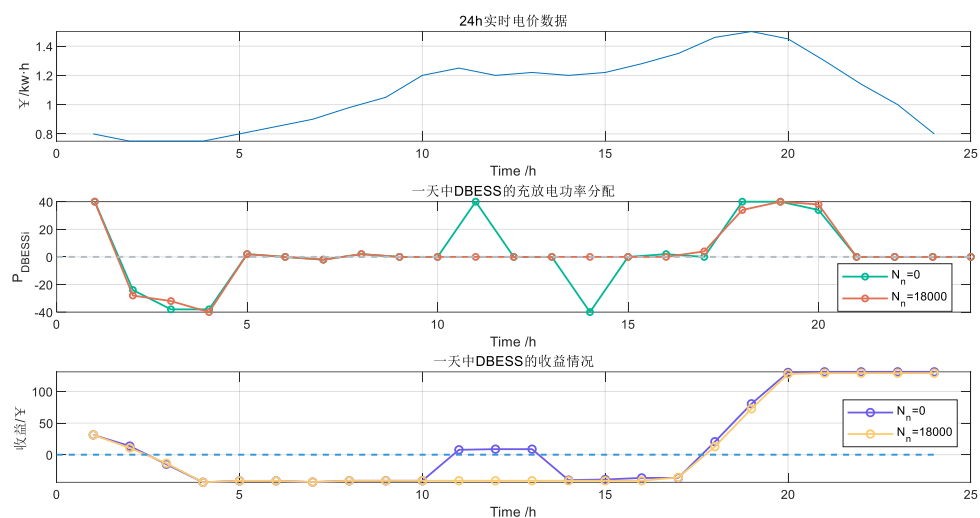
考虑老化成本，电价参数不变，DBESS 采用钛酸铁锂电池，对比不同老化程度的 DBESS 在 24 小时内的充放电策略。（并未考虑电池效率的衰减，因为对于 24 小时的算例，电池效率变化不大，可以忽略）

##### 4.2.1 算例参数

参数类型	DBESS1	DBESS2
额定容量 $Q_N$	120 kW · h	额定容量 $Q_N$
额定充放电功率 P	40 kW	额定充放电功率 P
初始荷电状态 SOC	50%	初始荷电状态 SOC
初始效率 $\eta$	98%	初始效率 $\eta$
购置成本 $C_{i,ap}$	2500 × 120 ¥	2500 × 120 ¥
最大充放电次数 N	20000	20000
已充放电次数 $N_n$	0	18000
$K_p$	1e-4	1e-4
E	120 kW · h	120 kW · h

注：  $K_p$  和 E 是我猜的，  $K_p$  取得太大使得逐周期老化成本较高，以至于不同电池对  $N_n$  不敏感，都趋向于保守充放电。

4.2.2 算例结果



二者在总体趋势上都满足低储高发，容易看出，当 DBESS 已充放电次数较高时，在 24h 充放电策略上相比新电池更加保守。（5h~17h 这一段）

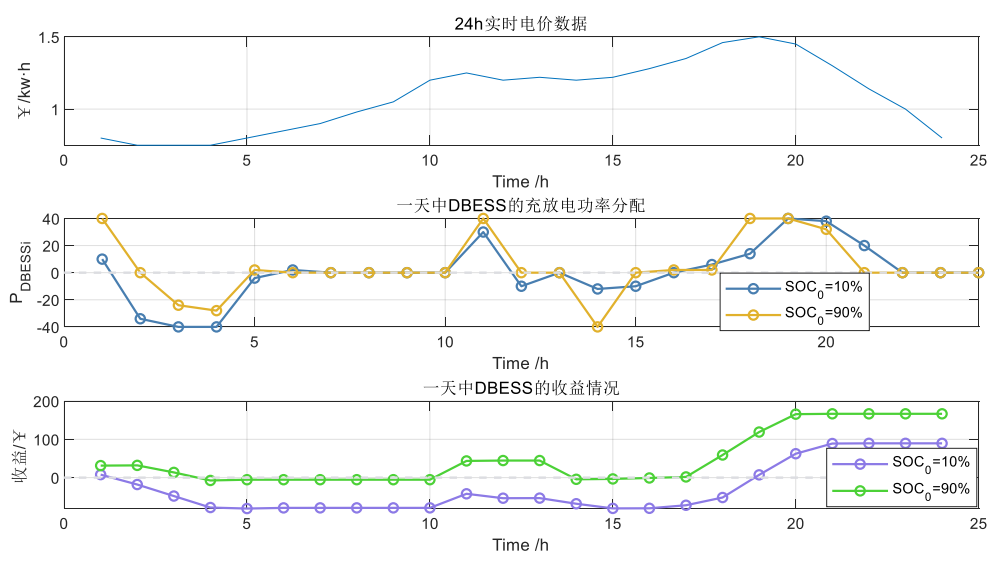
4.3 算例三 不同初始 SOC 的 DBESS 策略对比

沿用算例一的电价参数和 DBESS 参数，考虑老化成本，不考虑效率衰减。对比不同初始 SOC 状态的 DBESS 在 24h 内的充放电策略。

4.3.1 算例参数

初始 SOC	10%	90%
--------	-----	-----

4.3.2 算例结果



初始 SOC 状态为 90% 的 DBESS 一天的总体收益高于初始 SOC 为 10% 的 DBESS（达到 150 元以上），两种初始 SOC 状态的 DBESS 充放电策略趋势相同，但初始 SOC 状态为 90% 的 DBESS 放电功率较高。

#### 4.4 算例四 算法性能对比

沿用算例一中的数据，用强化学习与模拟退火(Simulated Annealing, SA)算法，对比 DBESS 的充放电策略和收益情况。由于算法运行时间与退火参数和强化学习参数设置相关，两种算法运行时间没有可比性。但是可以肯定的是，强化学习得到的 Q 矩阵，如果模型的泛化能力较强，则可以直接沿用作为后续的最优策略，节省了实时优化的计算资源。

##### 4.4.1 算例数据

模拟退火参数	
初始温度 $T_0$	500000000
温度衰减率 $\Delta T$	0.5%
终止条件	$T < 0.01$
搜索衰减比 $K$	0.65

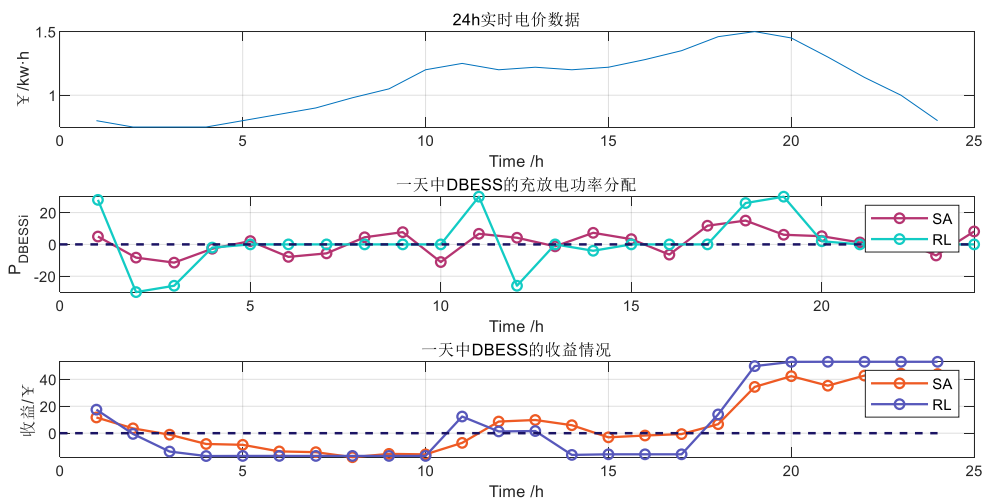
强化学习参数	
最大学习次数 $N_{max}$	30000
学习率 $\alpha$	1
奖励衰减率 $\lambda$	1

注：模拟退火的冷却模型为指数衰减，随机搜索模型自己设计，公式如下：

$$dP = ((-P_t - P_N) / T_0) + (2P_N / T_0)rand \cdot (T_0 + K(T - T_0)e^{-T})$$

此搜索模型可以满足一次性搜索满足定义域约束的状态变量，并且搜索衰减可控。其中  $P_t$  为  $t$  时刻的充放电功率， $P_N$  为额定充放电功率，rand 为 0~1 之间的随机数， $K$  为搜索衰减比，控制  $K$  可以控制搜索过程。增大  $K$  可以防止后期无法生成满足约束的解陷入死循环的情况，减小  $K$  可以使得最终解精确。

##### 4.4.2 算例结果

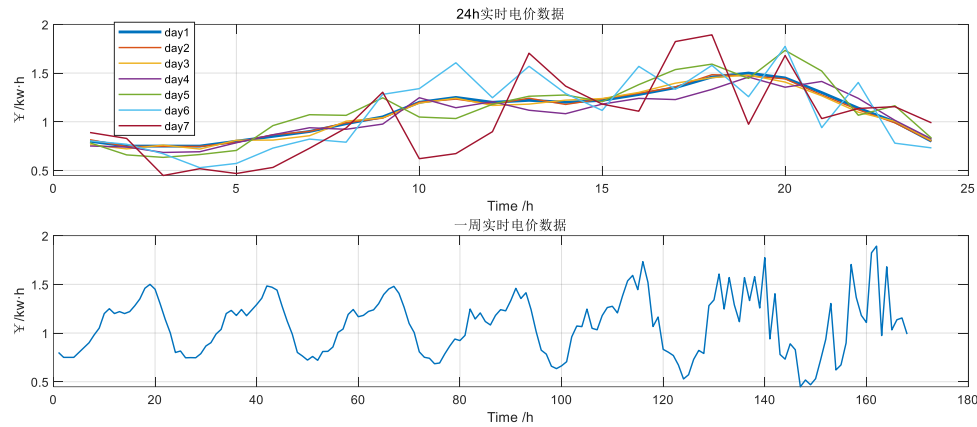


SA 的策略不能很好的体现出低储高发，有可能是因为搜索程度不够或者搜索算法不够好，收益上也低于 RL。(注：这里的收益是乘了 DBESS 低储高发抽成系数 0.8)

4.5 算例五 强化学习策略的泛化能力

将 24h 电价数据拓展到一周，第一天为原始的电价数据，第二天的实时电价数据在第一天基础上加入 $\pm 2\%$ 的扰动，第三天在第一天基础上加入 $\pm 5\%$ 的扰动...，使用算例二的 DBESS 数据和值函数矩阵  $Q$ ，每一天不进行重复学习的策略调整，而是直接沿用  $Q$  矩阵的策略适应新的电价数据。

4.5.1 算例数据



扰动程度						
第一天	第二天	第三天	第四天	第五天	第六天	第七天
0%	2%	5%	10%	20%	30%	50%

4.5.2 算例结果

