Joint Optimization of Thermostatically Controlled Appliance and Electric Vehicle Based on Adaptive Dynamic Programming

Yi Liu[1, 2, a], Yun Zhang[1, b, *], Sizhe Chen[1, c]

[1]School of Automation, Guangdong University of Technology, Guangzhou 510006, China

[2] School of Information Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China

[a]email: baddielao@126.com, [b]email: yun@gdut.edu.cn, [c]email: cszscut@126.com

Keywords: Home Energy Management; Adaptive Dynamic Programming; Optimal Control; Time-varying; Thermostatically Controlled Appliance; Electric Vehicle.

Abstract. There are time-varying parameters within the performance index function and dynamic model of home energy optimization (HEO) which contains thermostatically controlled appliances (TCAs) and electric vehicles (EVs). And TCAs and EVs may be off-line. Adaptive dynamic programming (ADP) can not be applied directly to the optimization. In this paper, these problems are solved by transformation of the performance index function and dynamic model, and ADP is applied to the optimization with guarantees of convergence and optimality. The performances of the transformation are given in numerical results.

Introduction

The energy consumption of TCAs, including heating, ventilation, and air-conditioning (HVAC), exceeds half of the household energy consumption[1]. With the popularity of EVs, EVs become important storage devices in smart grid. HEO with HVACs and EVs is a hot research topic in the literature [1-5].

HEO is an infinite-horizon problem. The cost of every time step in the future should be considered. However, as the dimension of solution increases while the number of time step increases, intelligent algorithms such as particle swarm optimization confront more difficulty to obtain the optimal solution. On the other hand, dynamic programming (DP) solves the Hamilton-Jacobi-Bellman (HJB) equation accurately and gives the optimal solution exactly [6]. But because of the "curse of dimensionality" [6], DP meets great burdens of computation and storage while the dimension of solution increases. Therefore, HEO is generally simplified as finite-horizon problem. The time horizon is usually set as one day and the number of time step is usually set as 24 [1-5].

ADP is an iterative algorithm[7]. It avoids the "curse of dimensionality", and it can solve infinite-horizon dynamic problem. The convergence and optimality of ADP without time-varying parameters were proven in [8-12], which assumed the systems are always on-line. Reference [13] assumed the dynamic price and uncontrollable loads were periodic time-varying parameters, applied ADP to schedule the power of battery and proved the convergence and optimality. But the battery was always on-line and its model was a time-invariant function. However, Thermal system is influenced by outdoor temperature, solar irradiance, and so on. System of EV is influenced by user's demand of travel as well. These systems are time-varying. Moreover, These systems will be off-line while they are shut down by user. Therefore, ADP can not directly applied to solve HEO with HVACs and EVs.

In this paper, the off-line problem are solved, the time-varying parameters in system's model and performance index function are eliminated through an assumption, and simulation results show the performances of the assumption.

Problem Formulation

The HEO system in the paper includes a controller, HVACs, EVs, and other home appliances. The HVACs and the EVs are controlled by the controller. The other home appliances are uncontrollable.

Let Δt be the time interval and k be the time step index, $k = 0, 1, 2, \mathbf{L}$. E_k is the battery's energy of EV in time step k and P_k^{EV} is the power of EV in k. P_k^{EV} is positive when EV charges and negative when discharges. Define a variable as

$$EV_k^p = \begin{cases} 1, & \text{EV is parked at home in time step } k \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

which denotes whether EV is parked at home in time step k. Based on Reference [1], EV's model is expressed as

$$E_{k+1} = E_k + EV_k^p c(P_k^{EV}) P_k^{EV} \Delta t - (1 - EV_k^p) E^{lost} \tag{2}$$

where $c(P_k^{EV})$ is the efficiency of EV's battery and E^{lost} is the energy consumption of EV in each time interval while EV is not at home. The second term of Eq. 2 shows that EV could be charged or discharged only when it is parked at home. The third term denotes that EV consumes energy when it is not at home.

Indoor temperature of a home is influenced by cooking, solar irradiance, wind speed, humidity, outdoor temperature, HVAC power, etc. For simplicity, we assume solar irradiance, outdoor temperature, and HVAC power are the major factors of thermal model [1]. Let T_k^{in} be the indoor temperature in time step k, T_k^{out} be the outdoor temperature in k, P_k^{HVAC} be the input power of HVAC in k, and f_k be the solar irradiance in k. Thermal model is described as

$$T_{k+1}^{in} = T_k^{in} + l_1(T_k^{out} - T_k^{in}) + l_2 HVAC_k^w P_k^{HVAC} + l_3 f_k \tag{3}$$

where l_1, l_2, and l_3 are coefficients, and

$$HVAC_k^w = \begin{cases} 1, & \text{HVAC is working in time step } k \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

denotes whether HVAC is working or not in k. It is zero when HVAC is turned off by user. Let

$$EV_k^d = \begin{cases} 1, & EV_{k-1}^p = 1 \text{ and } EV_k^p = 0 \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

be the variable that denotes whether k is the time step when EV just departs. P_k^H is defined as the total power consumption of the other home appliances in k, b_k is defined as the dynamic price in k. E^{max} and E^{min} are the maximum and minimum battery's energies of EV, $P^{EV,max}$ and $P^{EV,min}$ are the maximum and minimum powers of EV, $P^{HVAC,max}$ and $P^{HVAC,min}$ are the maximum and minimum input powers of HVAC. Let T_k^d be the desired temperature in k and E^d be the desired battery's energy of EV when EV just departs. T_k^d and E^d are set by user. Let N and M be the numbers of EV and HVAC, n and m be the indexes of EV and HVAC. The infinite-horizon performance index function which is expected to be minimized is described as

$$\hat{J} = \sum_{t=k}^{\infty} r^{t-k} \left(\hat{U}_1 + \hat{U}_2 + \hat{U}_3 + \hat{U}_4 \right)$$

$$\hat{U}_1 = c_1 \left[b_t \left(\sum_{n=1}^{N} EV_{t,n}^{p} P_{t,n}^{EV} + \sum_{m=1}^{M} HVAC_{t,m}^{w} P_{t,m}^{HVAC} + P_t^{H} \right) \right]^2$$

$$\hat{U}_2 = \sum_{m=1}^{M} HVAC_{t,m}^{w} \left(c_{2,m} (P_{t,m}^{HVAC} - \bar{P}_m^{HVAC})^2 + c_{3,m} (T_{t,m}^{in} - T_{t,m}^{d})^2 \right) \quad (6)$$

$$\hat{U}_3 = \sum_{n=1}^{N} EV_{t,n}^{p} \left(c_{4,n} (P_{t,n}^{EV} - \bar{P}_n^{EV})^2 + c_{5,n} (E_{t,n} - \bar{E}_n)^2 \right)$$

$$\hat{U}_4 = \sum_{n=1}^{N} EV_{t,n}^{d} c_{6,n} (E_{t,n} - E_n^{d})^2$$

where $r \in (0,1)$ is discount factor, c_1, $c_{2,m}$, $c_{3,m}$, $c_{4,n}$, $c_{5,n}$, and $c_{6,n}$ are weights, $\bar{P}^{HVAC} = 0.5(P^{HVAC,\max} + P^{HVAC,\min})$, $\bar{P}^{EV} = 0.5(P^{EV,\max} + P^{EV,\min})$, $\bar{E} = 0.5(E^{\max} + E^{\min})$. The utility function is composed of $\hat{U}_1$, $\hat{U}_2$, $\hat{U}_3$, and $\hat{U}_4$. $\hat{U}_1$ concerns the total electricity cost. In this paper, the battery of EV is not allowed to feed power to the power grid. $P_{t,n}^{EV}$ is ineffective for $\hat{U}_1$ when EV n is not at home. $P_{t,m}^{HVAC}$ is ineffective for $\hat{U}_1$ when HVAC m is turned off. The first term of $\hat{U}_2$ aims to make $P_{t,m}^{HVAC}$ close to the middle of its limits. It avoids $P_{t,m}^{HVAC}$ exceeds its limits. The second term of $\hat{U}_2$ aims to make $T_{t,m}^{in}$ close to the desired temperature. When HVAC m is turned off, it is ineffective for $\hat{U}_2$. The first term of $\hat{U}_3$ prevents $P_{t,n}^{EV}$ exceeds its limits, and the second term of $\hat{U}_3$ prevents $E_{t,n}$ exceeds the storage limits. EV n is ineffective for $\hat{U}_3$ when it is not at home. $\hat{U}_4$ aims to make $E_{t,n}$ close to the desired energy when EV n just departs. We define

$$x_k = [E_{k,1}, \mathbf{L}, E_{k,N}, T_{k,1}^{in}, \mathbf{L}, T_{k,M}^{in}]^T$$
$$u_k = [P_{k,1}^{EV}, \mathbf{L}, P_{k,N}^{EV}, P_{k,1}^{HVAC}, \mathbf{L}, P_{k,M}^{HVAC}]^T$$
$$l_k = [EV_{k,1}^{p}, \mathbf{L}, EV_{k,N}^{p}, HVAC_{k,1}^{w}, \mathbf{L}, HVAC_{k,M}^{w}, T_k^{out}, f_k]^T \quad (7)$$
$$r_k = [b_k, P_k^{H}, EV_{k,1}^{p}, \mathbf{L}, EV_{k,N}^{p}, HVAC_{k,1}^{w}, \mathbf{L}, HVAC_{k,M}^{w},$$
$$\qquad T_{k,1}^{d}, \mathbf{L}, T_{k,M}^{d}]^T$$

where x_k is the vector of state, u_k is the vector of control, l_k and r_k are the vectors of time-varying parameter. The problem described by Eq. 2, Eq. 3, and Eq. 6 could be conveniently represented as

$$\hat{J} = \sum_{t=k}^{\infty} r^{t-k} \hat{U}(x_t, u_t, r_t) \quad (8)$$

$$x_{k+1} = f(x_k, u_k, l_k). \quad (9)$$

As EV_k^{d} is determined by EV_k^{p}, it is not included in r_k. Eq. 6 shows that when EV n is not at home, $P_{t,n}^{EV}$ is arbitrary as it doesn't affect Eq. 6. And so do $P_{t,m}^{HVAC}$ when HVAC m is not working. Actually $P_{t,n}^{EV}$ must be equal to zero when EV n is not at home, and $P_{t,m}^{HVAC}$ must be zero when HVAC m is turned off. Therefore we define

$$\Lambda_k = diag(EV_{k,1}^{p}, \mathbf{L}, EV_{k,N}^{p}, HVAC_{k,1}^{w}, \mathbf{L}, HVAC_{k,M}^{w}) u_k \quad (10)$$

as the actual vector of control from the controller to the EVs and the HVACs.

ADP Algorithm

A discrete-time nonlinear system is described as

$$x_{k+1} = F(x_k, u_k).$$ (11)

where $x_k \in \Box^n$ is state vector, $u_k \in \Box^m$ is control vector, and k is the index of time step, $k = 0, 1, 2, \mathbf{L}$. $\forall x_k, u_k$, $F(x_k, u_k)$ is continuous. It is assume that Eq. 11 is under state feedback control, that is $u_k = u(x_k)$. The performance index function which is minimized by u_k is

$$J(x_k) = \sum_{t=k}^{\infty} g^{t-k} U(x_t, u_t)$$ (12)

where $g \in (0,1)$ is discount factor and $U(x_t, u_t) \geq 0$ is utility function. Eq. 12 can be written as

$$J(x_k) = U(x_k, u_k) + g J(x_{k+1}) = U(x_k, u_k) + g J(F(x_k, u_k)).$$ (13)

ADP algorithm is computed as follows

$$u_i(x_k) = \arg \min_{u_k} \{ U(x_k, u_k) + g J_i(F(x_k, u_k)) \}$$ (14)

$$J_{i+1}(x_k) = \min_{u_k} \{ U(x_k, u_k) + g J_i(F(x_k, u_k)) \}$$ (15)

where $u_i(x_k)$ is the iterative control law and $J_i(x_k)$ is the iterative performance index function. $i = 0, 1, 2, \mathbf{L}$ is the iteration step index. While the initial iterative performance index function $J_0(x) = 0$, the proofs of convergence and optimality have been presented [11]. That is, $\lim_{i \to \infty} J_i(x_k) = J^*(x_k)$ and $\lim_{i \to \infty} u_i(x_k) = u^*(x_k)$, where $J^*(x_k)$ is the optimal performance index function and $u^*(x_k)$ is the optimal control law.

Elimination of Time-varying Parameters

There is not time-varying parameter in Eq. 12. Therefore, the time-varying parameters within Eq. 8 and Eq. 9 must be eliminated before applying ADP to the optimization. The time-varying parameters have a strong periodicity in HEO. In this paper, the time-varying parameters are assumed as daily periodic functions of time. For simplification of the following derivation, we assume the time interval is one hour and there are 24 time steps in a day. Therefore the assumption is

$$l_j = l_{j+24}, \ r_j = r_{j+24}, \ j = k, k+1, \mathbf{L}.$$ (16)

l_k and r_k are known in the assumption, they are not the input variables of Eq. 8 and Eq. 9. Define

$$\hat{U}_t(x_t, u_t) = \hat{U}(x_t, u_t, r_t) \text{ and } f_k(x_k, u_k) = f(x_k, u_k, l_k).$$ (17)

Based on the assumption of Eq. 16, the problem described by Eq. 8 and Eq. 9 is rewritten as

$$\hat{J}(x_k) = \sum_{t=k}^{\infty} r^{t-k} \hat{U}_t(x_t, u_t)$$ (18)

$$x_{k+1} = f_k(x_k, u_k),$$ (19)

and we have

$$\hat{U}_j(\cdot) = \hat{U}_{j+24}(\cdot), \ f_j(\cdot) = f_{j+24}(\cdot), \ j = k, k+1, \mathbf{L}.$$ (20)

According to Eq. 19, Eq. 18 can be written as

$$\hat{J}(x_k) = \hat{U}_k\left(x_k, u_k\right) + \sum_{t=k+1}^{k+23} r^{t-k}\hat{U}_t\left(f_{t-1}(f_{t-2}(\mathbf{L}\left(f_k(x_k, u_k), u_{k+1}\right), \mathbf{L}), u_t\right)$$

$$+ r^{24}\hat{U}_{k+24}\left(x_{k+24}, u_{k+24}\right) + \sum_{t=k+25}^{k+47} r^{t-k}\hat{U}_t\left(f_{t-1}(f_{t-2}(\mathbf{L}\left(f_{k+24}(x_{k+24}, u_{k+24}), u_{k+25}\right), \mathbf{L}), u_t\right) \qquad . \tag{21}$$

$$+ r^{48}\hat{U}_{k+48}\left(x_{k+48}, u_{k+48}\right) + \sum_{t=k+49}^{k+71} r^{t-k}\hat{U}_t\left(f_{t-1}(f_{t-2}(\mathbf{L}\left(f_{k+48}(x_{k+48}, u_{k+48}), u_{k+49}\right), \mathbf{L}), u_t\right) + \mathbf{L}$$

Based on Eq. 20, Eq. 21 is transformed into

$$\hat{J}(x_k) = \hat{U}_k\left(x_k, u_k\right) + \sum_{t=k+1}^{k+23} r^{t-k}\hat{U}_t\left(f_{t-1}(f_{t-2}(\mathbf{L}\left(f_k(x_k, u_k), u_{k+1}\right), \mathbf{L}), u_t\right)$$

$$+ r^{24}\hat{U}_k\left(x_{k+24}, u_{k+24}\right) + r^{24}\sum_{t=k+1}^{k+23} r^{t-k}\hat{U}_t\left(f_{t-1}(f_{t-2}(\mathbf{L}\left(f_k(x_{k+24}, u_{k+24}), u_{k+25}\right), \mathbf{L}), u_{t+24}\right) \qquad . \tag{22}$$

$$+ r^{48}\hat{U}_k\left(x_{k+48}, u_{k+48}\right) + r^{48}\sum_{t=k+1}^{k+23} r^{t-k}\hat{U}_t\left(f_{t-1}(f_{t-2}(\mathbf{L}\left(f_k(x_{k+48}, u_{k+48}), u_{k+49}\right), \mathbf{L}), u_{t+48}\right) + \mathbf{L}$$

Define a new control law as

$$v_k = \{u_k, u_{k+1}, \mathbf{L}, u_{k+23}\} \tag{23}$$

and a new utility function as

$$V_k(x_k, v_k) = \hat{U}_k\left(x_k, u_k\right) + \sum_{t=k+1}^{k+23} r^{t-k}\hat{U}_t\left(f_{t-1}(f_{t-2}(\mathbf{L}\left(f_k(x_k, u_k), u_{k+1}\right), \mathbf{L}), u_t\right). \tag{24}$$

The performance index function is finally rewritten as

$$\hat{J}(x_k) = \sum_{L=0}^{\infty} r^{24L}V_k(x_{k+24L}, v_{k+24L}). \tag{25}$$

Besides Eq. 25, a new dynamic model is needed. According to Eq. 19, we have

$$\begin{aligned}
x_{k+24} &= f_{k+23}(f_{k+22}(\mathbf{L}\left(f_k(x_k, u_k), u_{k+1}\right), \mathbf{L}), u_{k+23}) \\
x_{k+48} &= f_{k+47}(f_{k+46}(\mathbf{L}\left(f_{k+24}(x_{k+24}, u_{k+24}), u_{k+25}\right), \mathbf{L}), u_{k+47}) \\
x_{k+72} &= f_{k+71}(f_{k+70}(\mathbf{L}\left(f_{k+48}(x_{k+48}, u_{k+48}), u_{k+49}\right), \mathbf{L}), u_{k+71})
\end{aligned} \tag{26}$$

$$\mathbf{L}$$

Define

$$g_k(x_k, v_k) = f_{k+23}(f_{k+22}(\mathbf{L}\left(f_k(x_k, u_k), u_{k+1}\right), \mathbf{L}), u_{k+23}). \tag{27}$$

Based on Eq. 20, Eq. 26 can be rewritten as

$$x_{k+24(L+1)} = g_k(x_{k+24L}, v_{k+24L}). \tag{28}$$

The problem described by Eq. 8 and Eq. 9 is transformed into an equivalent problem described by Eq. 25 and Eq. 28 under the assumption of Eq. 16. When the first time step k of the problem is confirmed, $V_k(\cdot)$ and $g_k(\cdot)$ are time-invariant. Based on Eq. 28, Eq. 25 can also be written as

$$\hat{J}(x_k) = V_k(x_k, v_k) + r^{24}\hat{J}(x_{k+24}) = V_k(x_k, v_k) + r^{24}\hat{J}(g_k(x_k, v_k)). \tag{29}$$

Let $U(\cdot) = V_k(\cdot)$, $F(\cdot) = g_k(\cdot)$, and $g = r^{24}$. According to Eq. 13, Eq. 29 can be solved by ADP with the guarantees of convergence and optimality.

Numerical Results

Assumption of Eq. 16 is the key of the tranformation. Actually, the time-varying parameters of different days are not completely the same. In this section, an ADP application for HEO is shown to test the performances of the assumption. For simplicity, we assume that the numbers of HVAC and EV are all one in the example.

The time interval is set as 1 hour and the first time step is $k = 0$. We assume the EV has following parameters: its battery's capacity is 20kWh; its maximum power is 6kW and minimum power is -6kW; its initial battery's energy is 15kWh; it consumes 1.3kWh in each hour while it is not at home; its battery's efficiency is 1; the user sets 17kWh as the value of desired battery's energy when the EV just departs; the time step when the EV just departs is normal distribution with the mean of 7:00 and the standard deviation of 2 hours, and the time step when the EV just arrives is normal distribution with the mean of 14:00 and the standard deviation of 2 hours [14]. We assume the HVAC runs in cooling mode and the coefficients of thermal model are $l_1 = 0.5$, $l_2 = -2.5$, and $l_3 = 4$. The maximum input power of the HVAC is 4kW. The initial indoor temperature is 27.5°C. The time step when the HVAC is just turned off is normal distribution with the mean of 20:00 and the standard deviation of 1 hour, the time step when the HVAC is just turned on is normal distribution with the mean of 22:00 and the standard deviation of 1 hour. The coefficients of performance index function are set as follows: $r = 0.95$, $c_1 = 0.2$, $c_{2,1} = 0.001$, $c_{3,1} = 0.001$, $c_{4,1} = 0.001$, $c_{5,1} = 0.001$, and $c_{6,1} = 0.02$. The average time-varying parameters of history are shown in Fig. 1(a). The actual time-varying parameters of the next 5 days are shown in Fig. 1(b). The time-varying parameters of different days are similar but not the same.

We assume the time-varying parameters of the next 5 days are predicted perfectly. We set the time horizon as 5 days and optimize based on DP. The results are the standards of following comparisons. In the following discussion, we call them the standard, which is shown in Fig. 2. The time horizon of the standard should be set as infinity, but it is finite here because of the limitation of simulation.

We analyze the results of the standard. The actual power of EV and the battery's energy of EV do not exceed their limits. The actual input power of HVAC is neither too large nor negative. The indoor temperature is close to the desired temperature when the HVAC is working. When the dynamic price and the total power consumption of the other home appliances are low, the EV charges. The battery's energy satisfies the desired one when the EV just departs. When the EV is not at home, the actual power of EV is zero and the battery's energy of EV is decreasing. The EV discharges when the dynamic price and the total power consumption of the other home appliances are relatively high, but almost stops charging or discharging when the dynamic price is low but the total power consumption of the other home appliances is high. The indoor temperature almost satisfies the desired one from 0:00 to 7:00 per day when the dynamic price, the total power consumption of the other home appliances, the outdoor temperature, and the solar irradiance are very low. The actual input power of HVAC increases greatly at noon because the outdoor temperature and the solar irradiance are high. But due to the impact of electricity cost, the indoor temperature isn't very close to the desired one. When the HVAC is turned off, the actual input power of HVAC is zero and the indoor temperature increases.

The following simulations are divided into 2 cases.

Case 1: the time-varying parameters of the next days can not be predicted. We assume the time-varying parameters of the next days are the same as the ones in Fig. 1(a). We apply ADP to the optimization. As mentioned in Introduction, the time horizon is generally set as 1day. We transform the problem into 5 new problems which time horizon is 1day and optimize based on DP. They are called the simplified new problems based on DP (SNP-DP). The results of case 1 are shown in Fig. 2(a). As the time-varying parameters of the next days are not accurate, the results of ADP and SNP-DP are not the same as the ones of the standard. For example, the EV leaves earlier than usual on the second day and the forth day, the battery's energy of EV can not reach the desired one. Furthermore, the difference between the results of SNP-DP and the standard is more obvious than the difference between the

results of ADP and the standard around 00:00 everyday as the cost after the end of time horizon is not considered in SNP-DP.

Case 2: the time-varying parameters of the next day are predicted accurately. We apply SNP-DP to the optimization. We assume the time-varying parameters of the next days are the same as the ones of the next day and apply ADP to the optimization. The problem is resolved again by ADP and SNP-DP with the latest predictions. The results of example 1 are shown in Fig. 11. As the time-varying parameters are predicted perfectly, the results of ADP, the standard and SNP-DP are overlap almost completely. The results of ADP and the standard are difference around 120:00, as the cost after the next 5 days is not considered in the standard. Therefore, the values of performance index function of ADP and the standard are almost the same in Table II. As analyzed in case 1, the results of SNP-DP are different with the ones of ADP and the standard around 00:00 everyday, and the values of performance index function of SNP-DP are the greatest ones.
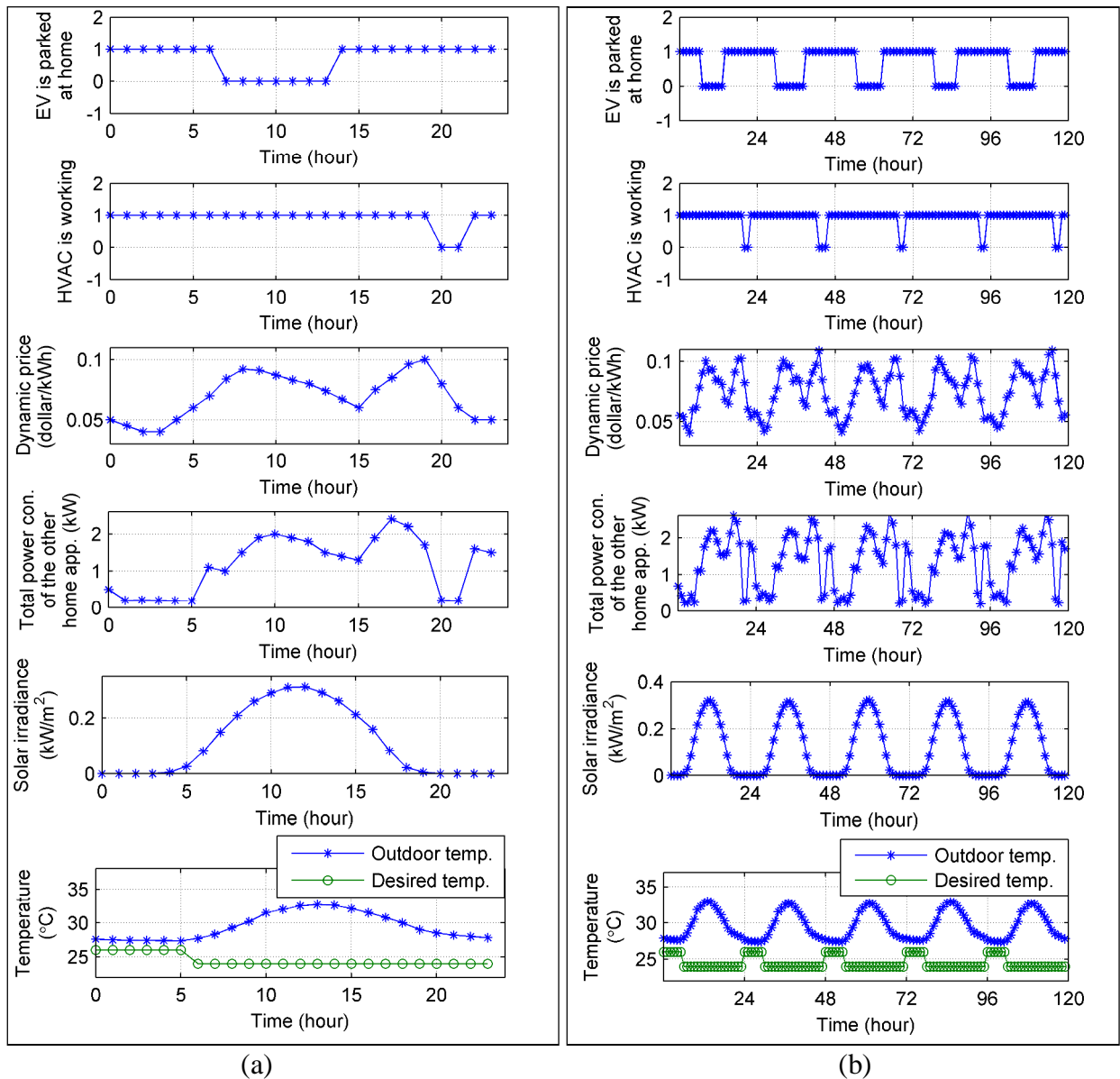

Fig. 1. Time-varying parameters. (a) Average time-varying parameters of history. (b)Actual time-varying parameters of the next 5 days.
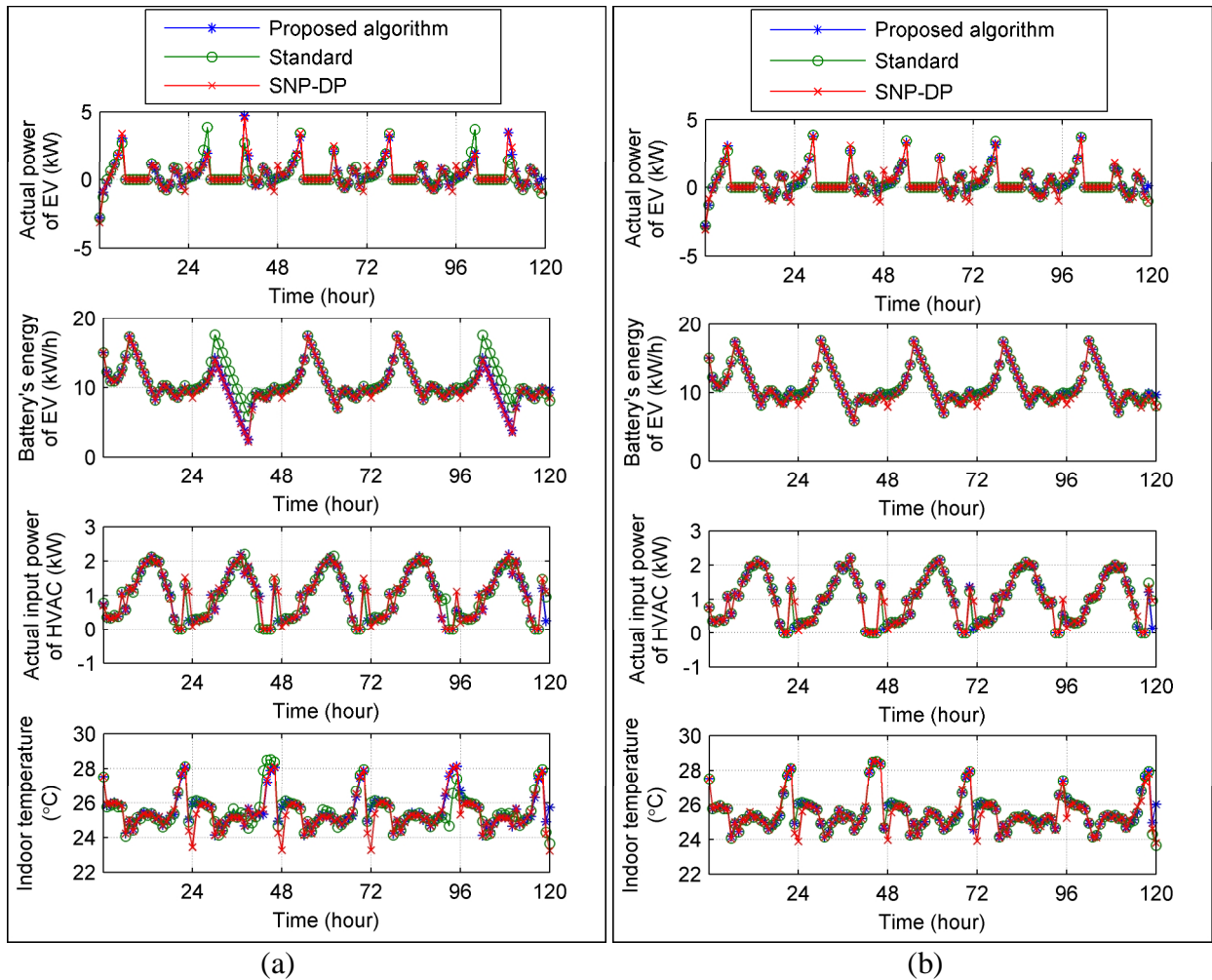
Fig. 2. Results of case 1 and case 2. (a) Results of case 1. (b) Results of case 2.

Conclusions

According to the periodic characteristic, the time-varying parameters in HEO are eliminated by transformations in this paper and ADP is applied to the HEO with HVACs and EVs. As future study, it is suggested to look into the problem of transformation while uncertain variables such as wind speed are included in HEO.

Acknowledgements

References

[1] D. T. Nguyen and L. B. Le: Joint optimization of electric vehicle and home energy scheduling considering user comfort preference. IEEE Trans. on Smart Grid, vol. 5 (2014), no. 1, pp. 188-199.

[2] M. H. K. Tushar, C. Assi, M. Maier, and M. F. Uddin: Smart microgrids: optimal joint scheduling for electric vehicles and home appliances. IEEE Trans. on Smart Grid, vol. 5 (2014), no. 1, pp. 239-250.

[3] M. Ghofrani, A. Arabali, M. Etezadi-Amoli, and M. S. Fadali: Smart scheduling and cost-benefit analysis of grid-enabled electric vehicles for wind power integration. IEEE Trans. on Smart Grid, vol. 5 (2014), no. 5, pp. 2306-2313.

[4] S. Xie, W. Zhong, K. Xie, R. Yu, and Y. Zhang: Fair Energy Scheduling for Vehicle-to-Grid Networks Using Adaptive Dynamic Programming. IEEE Trans. on Neural Networks and Learning Systems, vol. 27 (2016), no. 8, pp. 1697-1707.

[5] Y. Liu, Y. Zhang, K. Chen, S. Z. Chen, and B. Tang: Equivalence of multi-time scale optimization for home energy management considering user discomfort preference. IEEE Trans. on Smart Grid, DOI: 10.1109/TSG.2015.2510222 (2016).

[6] R. E. Bellman, in: Dynamic Programming. Princeton, NJ, USA: Princeton Univ. Press (1957).

[7] P. J. Werbos, in: Approximate dynamic programming for real-time control and neural modeling. in Handbook of Intelligent Control, D. A. White and D. A. Sofge, Eds. New York: Van Nostrand Reinhold (1992).

[8] J. J. Murray, C. J. Cox, G. G. Lendaris, and R. Saeks: Adaptive dynamic programming. IEEE Trans. on Systems, Man, and Cybernetics, Part C, vol. 32 (2002), no. 2, pp. 140-153.

[9] D. Liu and Q. Wei: Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems. IEEE Trans. on Neural Networks and Learning Systems, vol. 25 (2014), no. 3, pp. 621-634.

[10] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf: Discrete-time nonlinear HJB solution using approximate dynamic programming: convergence proof. IEEE Trans. on Systems, Man, and Cybernetics, Part B, vol. 38 (2008), no. 4, pp. 943–949.

[11] D. Wang, D. Liu, Q. Wei, D. Zhao, and N. Jin: Optimal control of unknown nonaffine nonlinear discrete-time systems based on adaptive dynamic programming. Automatica, vol. 48 (2012), no. 8, pp. 1825–1832.

[12] Q. Wei and D. Liu: A novel iterative θ-adaptive dynamic programming for discrete-time nonlinear systems. IEEE Trans. on Automation Science and Engineering, vol. 11 (2014), no. 4, pp. 1176-1190.

[13] Q. Wei, D. Liu, and G. Shi: A novel dual iterative Q-learning method for optimal battery management in smart residential environments. IEEE Trans. on Industrial Electronics, vol. 62 (2015), no. 4, pp. 2509-2518.

[14] C. Jin, J. Tang, and P. Ghosh: Optimizing electric vehicle charging with energy storage in the electricity market. IEEE Trans. on Smart Grid, vol. 4 (2013), no. 1, pp. 311–320.