13th IEA Heat Pump Conference
April 26-29, 2021 Jeju, Korea

Application of a deep reinforcement learning algorithm in household inverter air-conditioner temperature control

Myung Sup YOON[a,*], Won Sik YOON[a], Jong-Seok LEE[b]

[a]Energy Technology Center, Korea Testing Laboratory, 87 Digital-ro 26-gil, Guro-gu, Seoul, 08389, Republic of Korea
[b]School of Integrated Technology, Yonsei Unviersity, 85 Songdogwahak-ro, Yeonsu-gu, Incheon, 21983, Republic of Korea

Abstract

A deep reinforcement machine learning algorithm was applied to household inverter air-conditioner precision temperature control. Generally, air-conditioner temperature control aspects rely on the specific technology of the product, cooling room area size, outdoor temperature variation, and indoor building load variation when the set temperature is fixed. In this study, we fixed the test product and room size, and used the given variations of outdoor temperature and indoor building load over the course of one day. Even though the test product showed satisfactory performance at the remote controller set temperature of 26°C without a machine learning algorithm, we experimented with deep reinforcement learning performances to check whether the test product could follow general product performances or surpass the product ability in the precision temperature control by applying only high and low set temperatures alternatively. Training started with no disturbances of constant building cooling load and outdoor temperature. It resulted in reasonably accurate temperature control with the real training environments of Korean and Middle Eastern climates.

Keywords: Air-conditioner; Smart home; Deep reinforcement learning; Machine learning; Temperature control; AI

1. Introduction

Smart home appliances are a rapidly expanding industry. Korea is one of the leading countries in the smart home market, especially in the home appliance division. Premium air-conditioner and refrigerator products in Korea are now expected to embed smart-phone applications that can monitor energy consumption and be controlled from outside the house. However, there is further scope to improve them by adopting AI technology. Korean manufactured high-quality air-conditioners currently make it possible to perform target cooling by learning the occupant's positioning area and operating commands from the occupant's voice. But until now, these products seem to only borrow existing AI technology of vision and voice recognition from other fields.

Many HVAC (Heating, Ventilation, and Air-Conditioning) studies predict building cooling loads using a time series method and neural networks [1–4], fuzzy logic [5–7], and applying machine learning techniques in central air- conditioning systems [8–13]. However, it is not easy to find machine learning research in the field of household air-conditioners. For the energy usage comparison between inverter and constant speed air-conditioner technology, Yoon et al. [14,15] tested an air-conditioner under the actually varying temperature condition in an accurately controlled test chamber over a prolonged period. Referring to international standards such as ISO or EN, they suggested a one-point environmental test energy efficiency ratio (EER) or a more complicated multi-point interpolated energy efficiency ratio (seasonal EER) for the test of air-conditioner energy performance. These methodologies, however, cannot capture the dynamically improved performances of an AI driven air-conditioner because test chamber environments are statically fixed at each test point. Therefore, Yoon et al. [14,15] suggested a more realistic experiment. They simulated varying outdoor

* Corresponding author.
E-mail address: msyun95@ktl.re.kr

temperature and indoor building load conditions over the course of a day. This approach was analogous with a realistic car fuel economy test under real road load circumstances rather than a fixed high-speed condition.

In our study, a deep reinforcement learning algorithm was applied to a household inverter air-conditioner under an experimentally simulated varying environment (outdoor temperature, indoor building load). The purpose of this study was to check whether the machine learning algorithm could be applied to the household inverter air-conditioner temperature control and surpass the normal usage of the product.

2. Related work and background

Reinforcement learning (RL) is one of the most exciting and also one of the oldest areas of the machine learning discipline [16]. The problem of reinforcement learning is expressed in the Markov decision process (MDP), which is popular in the theory of dynamic programming [17]. Dynamic programming using value iteration or policy iteration evolved into the SARSA (State, Action, Reward, State, Action) on-policy problem, and later, the Q-learning (SARSA off-policy) [18].

A deep neural network (DNN) is an artificial neural network (ANN) consisting of several hidden layers between the input and output layers. One ANN study goes back to the Fukushima's convolution neural network (CNN) paper [19], and DNN is now used in various fields related to vision recognition [20,21]. In addition, language recognition studies use the modified DNN methods, which are RNN (Recursive Neural Network) or LSTM (Long Short-Term Memory) in machine learning [22–24].

In this study, we used the well-known DQN (Deep Q-learning Network) method that uses DNN for the Q-learning [18,25,26].

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_t + \gamma \cdot max_{a'} Q(S_{t+1}, a') - Q(S_t, A_t)] \quad (1)$$

where, (S_t, A_t) is the current state and action, (S_{t+1}, a') is the next state and action, $Q(S_t, A_t)$ is the quality of the agent in (S_t, A_t), R_t is the reward obtained by taking an action A_t from state S_t, and α and γ are the learning rate and the discount factor, respectively. In the iteration i, Q-learning updates by the following loss function

$$L_i(\theta_i) = \left(R_t + \gamma \cdot max_{a'} Q(S_{t+1}, a', \theta_{i-1}) - Q(S_t, A_t, \theta_i)\right)^2 \quad (2)$$

where, θ_i is the parameter of the Q-network at iteration i and θ_{i-1} is the network parameter from the previous iteration.

3. Methodology

3.1. Artificial neural network model

In this study, we used one input layer of 3 nodes, two hidden layers of 24 nodes each, and one output layer of 2 nodes. The input layer consists of three elements of state [$(T_{in}(t) - 26)$, dT_{in}/dt, $T_{dis}(t)$]. Each element represents indoor temperature (T_{in}), indoor temperature derivative (dT_{in}/dt), and air-conditioner discharge temperature ($T_{dis}(t)$) according to ISO 5151 [27] and ASHRAE 116 [28].

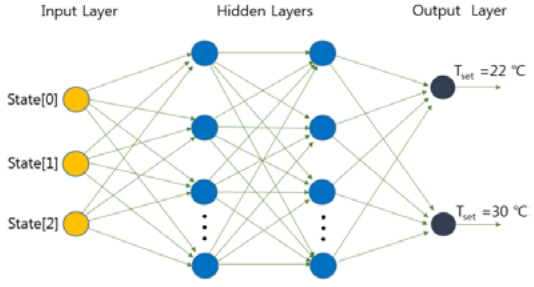

Fig. 1. DNN model for the inverter air-conditioner temperature control

The output layer gives action outputs to the remote controlled air-conditioner with the selected set temperature of each low and high product temperature [22, 30] °C. Even though the real test product has a wider controllable temperature range (16–31°C), we define action states here as two cases of set temperature to simplify the problem definition and reduce the calculation time when the target temperature is 26°C. These input, hidden, and output layers' artificial neural networks are generated by TensorFlow [29].

3.2. Simulation of the test chamber environment

Air-conditioner power consumption depends on indoor-side and outdoor-side environment conditions simultaneously. An indoor unit has an evaporator heat exchanger, and an outdoor unit usually has a condenser heat exchanger and a compressor component. If the outdoor-side temperature rises and the indoor cooling load is high, then the air-conditioner instantaneous power input usually increases. Therefore, international test standards such as ASHRAE and ISO request the exact fixing of the indoor and outdoor environment (temperature or humidity) for the rating test [28,29]. For example, the standard test condition (T1) of summer periods is (DBT/WBT) = 35°C/24°C and 27°C/19°C at each outdoor- and indoor-side chamber, respectively.

For the AI training of the product, we need time-varying real environment conditions rather than fixed-temperature conditions. Here, we assumed a summer situation of a typical region in Korea and the Middle East for the TRNSYS simulation [14,15]. The TRNSYS simulation program was used to obtain indoor building load $q_{build}(t)$ and outdoor temperature variation $T_{out}(t)$. The average temperature and building load variations during a summer's day can be obtained using TRNSYS climate data and analysis. The TRNSYS simulation condition is shown in Table 1.

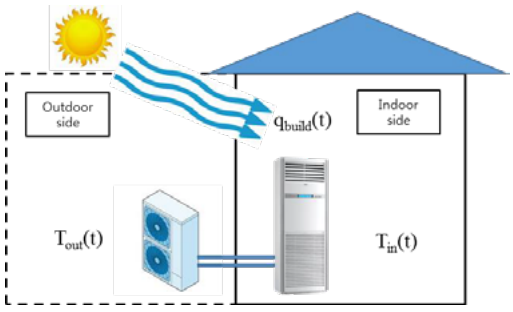

Fig. 2. Environment simulation concept (reproduced from Yoon et al. [14,15])

Table 1. TRNSYS simulation condition in Seoul & Middle East

Input variable	Simulation condition in Seoul	Simulation condition in M.E.
Space (single zone)	10m × 8m × 2.5m	10m × 8m × 2.5m
Window	60% of the southern wall area (15 m^2)	30% of the southern wall area (7.5 m^2)
Floor of space	Located between floors – up and down insulations	Located between floors – up and down insulations
Glass insulation	1.20 W/m^2K	1.720 W/m^2K
Wall insulation	0.21 W/m^2K	0.345 W/m^2K
Room temp. & humidity	26°C & 60%R.H.	26°C & 60%R.H.
Heating & light element	25 W/m^2	25 W/m^2
Persons	4 persons	4 persons
Infiltration	0.4 (1/h)	0.4 (1/h)
Ventilation	0.6 (1/h)	0.6 (1/h)

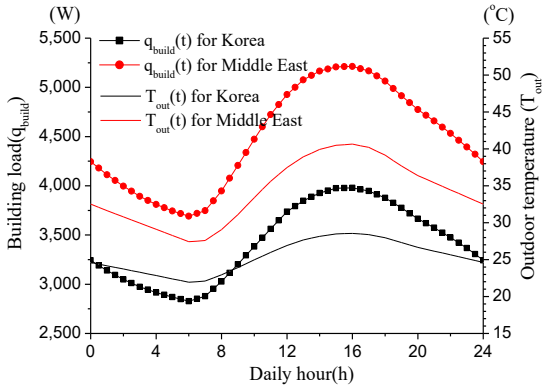

Fig. 3. Environment simulation result

3.3. Air-conditioner selection

A household room air-conditioner is categorized largely by constant speed compressor type and inverter compressor type. This study prepared an inverter compressor split wall mount air-conditioner. A constant speed air-conditioner's compressor turns off and on according to whether the room environment temperature satisfies the remote controller set temperature or not. In contrast, an inverter air-conditioner can control compressor speed to adjust the needed cooling capacity, which varies over time [14,15]. In this study, we simply control the set temperature by using an IR remote controller because each component's control method or signals (e.g., compressor speed, fan speed, expansion valve openness) are usually hidden by the manufacturer.

Even though the thermostat switching method of an inverter air-conditioner differs depending on the technology of each manufacturer, a dead band thermostat is usually used. This means that an air-conditioner starts to change the operation state when a certain activation range of temperature difference is satisfied between the measured room temperature and the product set temperature.

Table 2. Specifications of test air-conditioner

Contents	Name plate values
Air-conditioner type	Split wall mount
Compressor type	Inverter type
Electricity	Single phase, 220 V, 60 Hz
Rated cooling capacity	6 500 W
Rated power input	1 750 W
Max. power input	2 900 W
Min. power input	500 W
Refrigerant	R410a, 2kg

3.4. Deep reinforcement learning algorithm coding

The problem definition is:

- State : $[(T_{in}-26), dT_{in}/dt, T_{dis}]$
- Action : [22°C, 30°C]
- Reward : +1 pt, when T_{in} in the range of (26 ± 0.5)°C
- Penalty : –100 pts, when T_{in} out of (26 ± 0.5)°C

- Each episode ending : when penalty
- Environments during simulation time:
  - Day 1 : constant T_{out} = 28°C, q_{build} = 3,000 W
  - Day 2 : constant T_{out} = 28°C, q_{build} = 3,000 W (repeat one more day)
  - Day 3 : Korean summer climate, as Fig. 3
  - Day 4 : Middle Eastern summer climate, as Fig. 3
- Reset : T_{set} is reset within (26 ± 0.2)°C when each episode ends.

The coding flow is shown in Fig. 4. The basic algorithm is the same as established deep Q-learning, with experience replay [25], but differs by the addition of the training condition, T_{in} = (26 ± 0.5)°C, and reset condition T_{in} = (26 ± 0.2)°C.

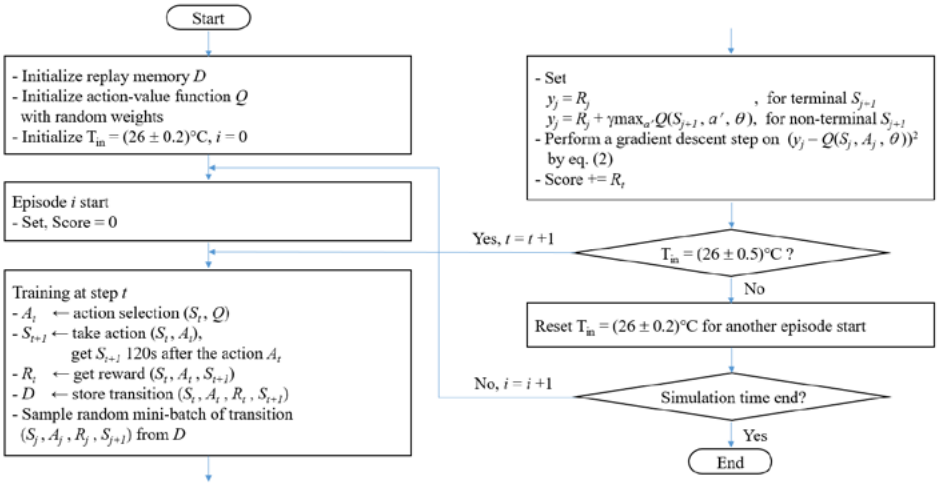

Fig. 4. Deep reinforcement learning coding flow

3.5. Air-conditioner remote control

Our deep reinforcement learning algorithm generates two cases of agent action, which are the remote controller setting temperatures of up and down. One of these action signals is transferred to the Arduino-based IR signal emitter. The IR signal changes the setting temperature of the air-conditioner. Before undertaking the IR operations, we need to record the infrared signals for each setting temperature of the remote controller. These Arduino IR signal-receiving and emission codes are well established in the Arduino-IRremote github library (https://github.com/z3t0/Arduino-IRremote). Figure 5 presents circuit diagrams of the IR signal receiver and emitter.

The block diagram for the experiment set-up is shown in Fig. 6. The test facility has one computer system for the facility environment simulation and data acquisition. The necessary data for the machine learning state inputs are transferred to a second PC. The second PC conducts a machine learning calculation and sends action output to the Arduino IR signal emitter.

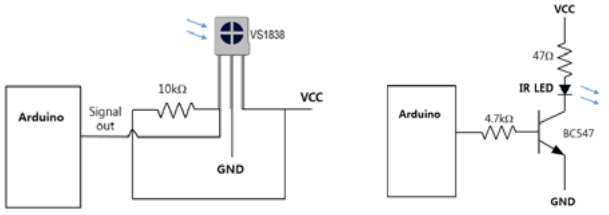

Fig. 5. IR signal receiver (left) and emitter (right) circuit

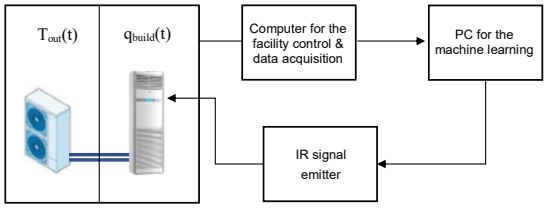

Fig. 6. Block diagram for the experiment set-up

4. Experiments and results

Even without AI, today's commercially available inverter air-conditioner technology is well established and can adjust building cooling demand to the desired room temperature. It offers satisfactory performance for the customer. However, this study aims to show that combining an inverter air-conditioner with RL machine learning can demonstrate better performance, especially in temperature precision control. First, normal usage of an air-conditioner with the fixed set temperature case, T_{set}=26°C, was experimented. Second, a simple thermostat algorithm was applied using computer control. Third, DQN RL training was started with a reward at T_{in}=(26±0.5)°C. Fourth, the same environment as the third case was maintained to check training reliability. Fifth, Korean summer climate environment variation was applied as a disturbance. Sixth, a Middle Eastern climate environment was applied as a harsh condition.

4.1. Without machine learning (normal usage of the product)

The recommended temperature setting for air-conditioners varies from country to country. In Korea, 26°C indoor room temperature is recommended. In order to compare the effect of RL machine learning, here we set the remote controller to 26°C without a machine learning algorithm and checked the indoor room temperature ($T_{in}(t)$) variation pattern under the Korean climate environment of outdoor temperature ($T_{out}(t)$) and building load ($q_{build}(t)$) change over the course of a day.

We introduced an inverter compressor-type air-conditioner to fine tune the indoor temperature, but the results in Fig. 7 are not very precisely controlled. Searching for other papers' results of different household inverter air-conditioner models does not change this fact [14,15]. The reason for this temperature control pattern is that ordinary HVAC products perform dead band control in order to reduce the switching frequency and increase the operating efficiency of the product [30]. During one day (24 h), as shown in Fig. 7, the maximum of T_{in} was 27.46 and the minimum was 25.00°C. Although the remote controller set temperature was fixed at 26°C, room temperature T_{in} oscillates with a deviation of approximately ±1°C.

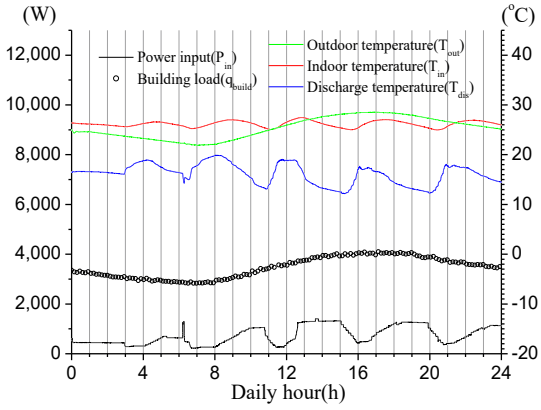

Fig. 7. Normal operation of the air-conditioner with T_{set} = 26°C

4.2. Simple computer thermostat algorithm control

In this case, the behavior of the product was identified by applying a simple thermostat algorithm rather than allowing the product to operate on its own. In Fig. 8(a), if the indoor temperature (T_{in}) was below 26°C, the product was remote controlled by an Arduino LED emitter at the value of 27°C. When T_{in} was over 26°C, the remote control value was set as 26°C. That is, a simple thermostat operation was implemented by computer algorithm. As another case of simple computer thermostat control, in Fig. 8(b), if T_{in} was below 26ºC, the product was remote controlled at the value of 30°C. When T_{in} was over 26°C, the remote control value was set as 22°C immediately.

Compared with Fig. 7, showing the normal operation under the T_{set} = 26°C, Fig. 8(a) of the simple thermostat algorithm shows similar graph patterns with some more fluctuations in indoor temperature (T_{in}), power input (P_{in}), and product discharge temperature (T_{dis}). Peaks in discharge temperature (T_{dis}) near 2, 6, and 10 hours were caused by the abrupt decrease of P_{in}. This abrupt change of P_{in} was due to the T_{set} signal of 27°C. To speculate on the control behavior of the product, we analyzed the point where the 26°C gray baseline meets T_{in}. At the 1.5 hour position, the room temperature (T_{in}) meets the downward 26°C line, and the automatic program changes the T_{set} set point to 27°C, but only slightly reduces power. In practice, the point at which power is drastically reduced is around 1.7 hours where the room temperature is at the local minimum. At 12 and 16 hours, T_{in} declines and meets the 26°C line. But, P_{in} does not decrease abruptly. For this reason, even though we cannot know the exact control design information of the original product, we can infer that the product uses dead band thermostat control.

Figure 8(b) shows the result when we use T_{set} of 22 and 30°C, instead of 26 and 27°C for the simple thermostat control. Because 22 and 30°C of T_{set} are far from our target of 26°C, immediate control is possible regardless of the dead band control of the product. The on and off graph line of the power input (P_{in}) and oscillating indoor temperature (T_{in}) are typical control aspects of a constant speed (or fixed speed) compressor air-conditioner. It is similar with the constant speed air-conditioner result (Fig. 3) of Yoon et al. [14].

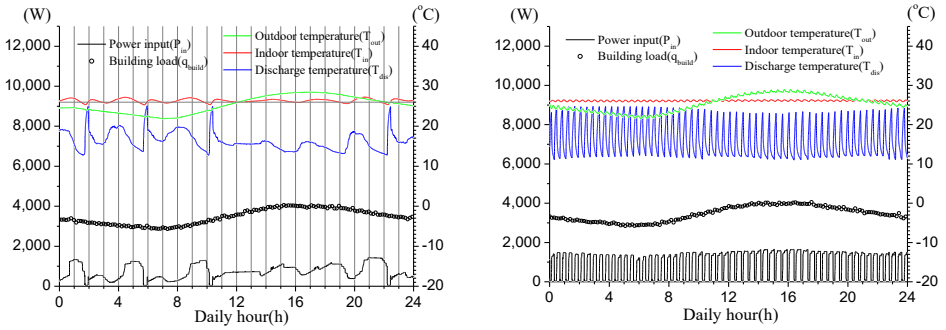

Fig. 8. Simple computer thermostat algorithm of (a) T_{set} = 26°C & 27°C (b) T_{set} = 22°C & 30°C

4.3. With machine learning algorithm (day 1 of RL)

On day 1 of reinforcement learning (RL), the test chamber environment was maintained at a constant outdoor temperature (T_{out}) of 28°C and a building load (q_{build}) of 3 000 W to reduce disturbances, as shown in Fig. 9. Even though the tested air-conditioner is an inverter type, which has typically small adjustment characteristics of power input (fine saw-tooth shape increase or decrease), such as in Figs. 7 and 8(a), in this case, Fig. 9 shows on and off switching characteristics as the indoor temperature was only controlled by the action of T_{set} high (30°C) and low (22°C).

Usually, pure computer RL machine learning such as in Atari games [25,26] takes less than a few hours. But real-world machine learning can take much longer, especially in the thermo-fluid system, which has a slow response time. Fortunately, this real-world air-conditioner RL study "only" took 16 hours to solve the problem. T_{in} in Fig. 9 is stabilized after 16 hours. The graph patterns of air-conditioner discharge temperature (T_{dis}) and power input (P_{in}) are irregular prior to 16 hours, so that sometimes the large time span of power-on duration can be seen (e.g., 9~10 h, 13 h). After 16 hours, P_{in} and T_{dis} show regular graph patterns.

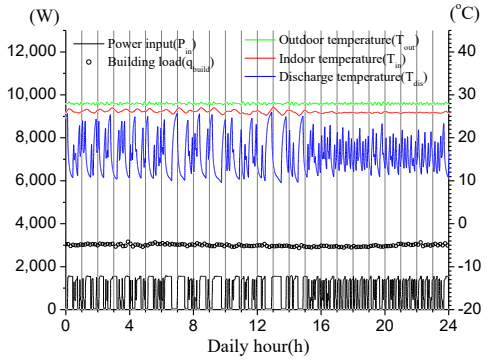

Fig. 9. Day 1 of RL result under the constant environment

An enlarged picture of the indoor temperature (T_{in}) is shown in Fig. 10. The red portions of the graph reflect the region where the reward +1 pt is taken. If T_{in} escapes the (26±0.5)°C range, each episode of training ends. For the restart of the training, the program algorithm forces the temperature down or up to reset T_{in} within (26±0.2)°C. The obtained scores at each episode are shown in Fig. 11. In this study, approximately one episode matches one hour before problem stabilization (16 h). It shows fluctuation from 2 to 25 scores. Even though Fig. 11 has no score information after 17 hours, episode 18 recorded 1,518 pts at the moment near the start time of day 4 (Fig. 15) when the environment changes abruptly to a harsh Middle Eastern climate condition.

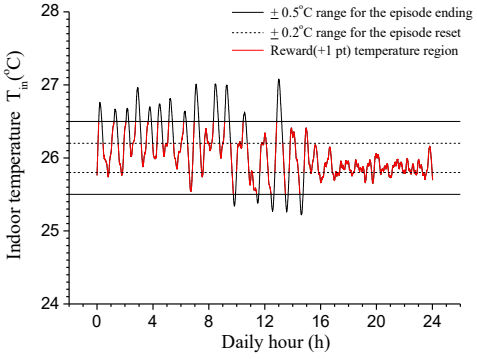

Fig. 10. Enlargement of the indoor temperature (T_{in}) graph during day 1

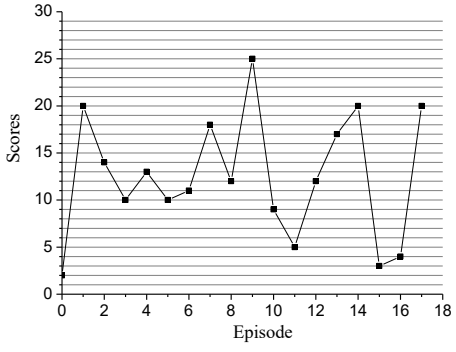

Fig. 11. Scores at each episodes during day 1

4.4. With machine learning algorithm (day 2 of RL)

On day 2 of RL, we repeated the same environment of constant T_{out} (28°C) and q_{build} (3 000 W) as day 1 of RL. Figure 12 shows a very well-trained and stabilized situation, and episode 18 continued during day 2 of RL. One thing to mention here is that throughout this work, there is a slight micro-vibration of the outdoor temperature (T_{out}) as a whole. This phenomenon is caused by the vibration of the condenser heat from the outdoor unit as the air-conditioner is turned on and off. Therefore, the fluctuation time matches between T_{out} and P_{in}. Normally, there are no such temperature interferences because air-conditioners are operated in a steady state on mode when products are tested according to ISO standard.

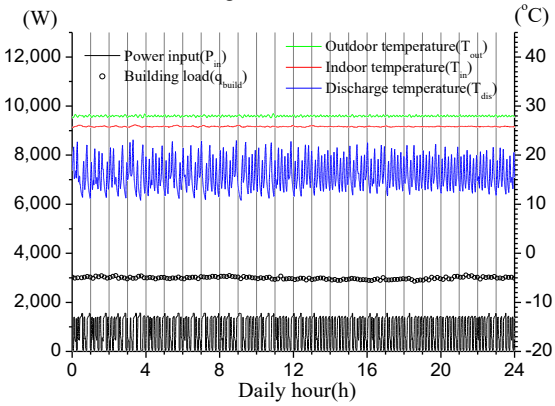

Fig. 12. 2nd day of RL result under the constant environment

4.5. With machine learning algorithm (day 3 of RL)

On day 3 of RL, a Korean summer climate (Fig. 3) was applied to the outdoor temperature ($T_{out}(t)$) and building cooling load ($q_{build}(t)$) as disturbance variations. Some other papers [2–4] have used these kinds of thermal variation as LSTM periodic state inputs, but in this study we did not use them as state inputs. Therefore, $T_{out}(t)$ and $q_{build}(t)$ variables act as unknown disturbances. From Fig. 13, $T_{out}(t)$ and $q_{build}(t)$ have minimum values near 6–7 a.m. and maximum values near 3–4 p.m. Indoor temperature (T_{in}) seems to converge well without change. However, the discharge temperature (T_{dis}) and power input (P_{in}) exhibit some different aspects with the constant environment case of Fig. 12. P_{in} near the maximum variation (3–4 p.m.) shows more on mode time span rather than minimum variation interval (6–7 a.m.). Also by the same reason, T_{dis} records a lower discharge temperature near 3–4 p.m. than 6–7 a.m. This means RL machine learning DNNs are adapted to the environment.

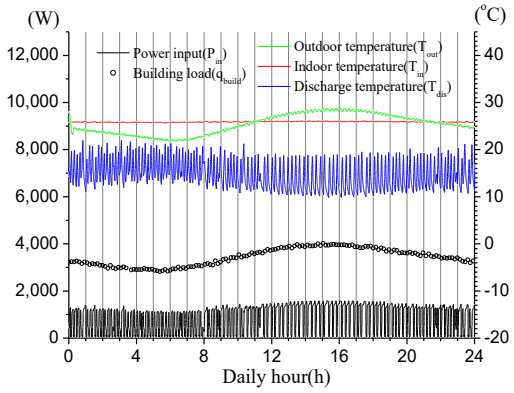

Fig. 13. Day 3 of RL result under Korean summer climate

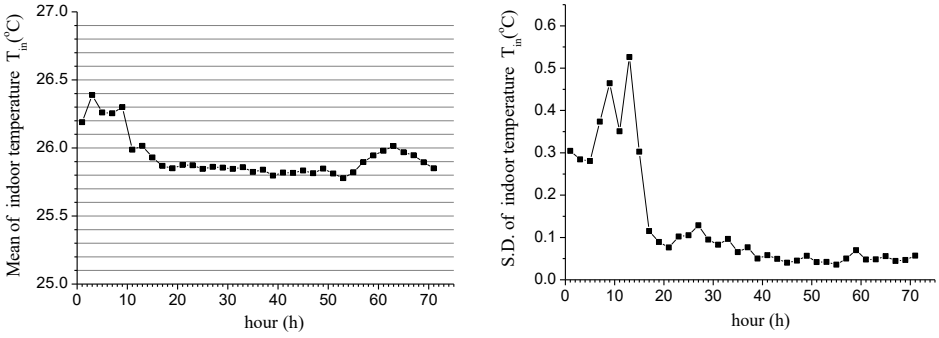

Fig. 14. Mean (left) and S.D. (right) of indoor temperature (T_{in}) during total three days at every 2 hour

The quality of T_{in} data during over the three days of RL was compared in Fig. 14. Because the training was stabilized after 16 hours of day 1, the mean value and standard deviation of T_{in} also show stabilization after 16 hours. However, the mean value shows one high peak at 64 hours (= 2 days and 16 h). This peak is due to environmental variations (T_{out} and q_{build}) during day 3 of RL. In any event, T_{in} exhibits a high accuracy when comparing the RL case of Figs. 12 and 13 with the normal usage (T_{set} = 26°C) case of Fig. 7.

4.6. With machine learning algorithm (day 4 of RL)

In this section, the environmental conditions of the Middle East were applied successively after the Korean environment, since summer temperatures in the Middle East are the harshest on the planet, which can act as big environmental disturbances in this study. In Fig. 15, because the start value of $q_{build}(t)$ abruptly changed from the Korean climate ending $q_{build}(t)$ value 3 250 W to the Middle East start value of 4 250 W, it shows disturbed T_{in} temperature near the start time. Therefore, we obtained a renewal of episode number by 19 at this moment. Episode 18 ended with the score of 1,518 pts at this position.

During the day time, there is a time interval in which power input (P_{in}) shows only an on mode without off mode. This means that the tested air-conditioner cooling capacity is insufficient to cover the building load (q_{build}). Notably from 13.5 hours to 19.8 hours, the indoor temperature (T_{in}) is over 26.5°C, even with the full operation of the product. Because the specification of the tested product (Table 2) is for moderate temperature condition (ISO T1 condition), the high temperature condition (ISO T3 condition) of the Middle East cannot be applied to this product. Nevertheless, this study conducted tests in the Middle East temperature condition to identify the limitations of RL machine learning due to harsh environmental conditions. Results show that even though the product capacity is insufficient, the trained deep neural network (DNN) continues to send 22°C of signal to cool the room.

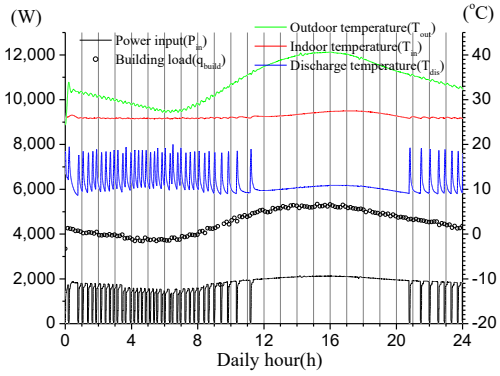

Fig. 15. Day 4 of RL results under the Middle Eastern summer climate

Table 3. Temperature mean, S.D., minimum, maximum values and power consumptions

Case	Mean of T_{in}	S.D. of T_{in}	Min. of T_{in}	Max. of T_{in}	P_{24h} (kWh)
Normal usage of T_{set} =26ºC (Fig. 7)	26.19	0.600	25.00	27.46	17.63
Simple thermostat of T_{set} =26ºC, 27ºC (Fig. 8(a))	26.39	0.431	25.34	27.29	17.52
Simple thermostat of T_{set} =22ºC, 30ºC (Fig. 8(b))	26.09	0.163	25.78	26.37	21.74
Day 1 of RL (Fig. 9), no climate disturbance	26.07	0.362	25.22	27.08	20.39
Day 2 of RL (Fig. 12), no climate disturbance	25.83	0.081	25.65	26.18	19.33
Day 3 of RL, Korean climate (Fig. 13)	25.90	0.088	25.66	26.12	20.23
Day 4 of RL, Middle Eastern climate (Fig. 15)	26.29	0.551	25.69	27.51	38.99

Finally, the temperature control qualities and power consumptions for each day are summarized in Table 3. The mean value and standard deviation of the T_{in} temperature show good results on day 2 and day 3 of RL. Similarly, the minimum and maximum deviations from the target temperature (26°C) also show minimum errors on day 2 and day 3 of RL. It shows that the application of DQN reinforcement learning in this study can surpass the temperature control ability of the product in normal usage or simple thermostat control cases. In terms of power consumption during one day (P_{24h}), this RL study case (about 20 kWh) is not energy efficient than normal usage case (17.63 kWh). This is because the product was machine-learned only using two action control temperatures of 22 and 30°C to avoid the dead band control zone inherent in the original air-conditioner product and to simplify the RL problem.

5. Conclusion

Reinforcement learning (RL) using a DQN (Deep Q-learning Network) method was applied to household inverter air-conditioner temperature precision control. Three state variables of $[(T_{in}(t) -26), dT_{in}/dt, T_{dis}(t)]$ and two actions of 22, 30°C remote control are used in the DQN model. The practical application of DQN RL to home appliances is seldom found in papers. In this study, after 16 hours of training, we obtained a stabilized indoor temperature close to 26°C. Environmental disturbances were applied during the training using Korean and Middle Eastern summer climates. The Korean climate conditions, which can be covered by product capabilities, showed a more stable and accurate temperature control pattern. Comparing with the normal usage result of T_{set}=26°C (T_{mean}=26.19°C, S.D.=0.600°C, T_{min}=25.00°C, T_{max}=27.46°C) or computer simple thermostat control of T_{set}=22 and 30°C (T_{mean}=26.09°C, S.D.=0.163°C, T_{min}=25.78°C, T_{max}=26.37°C), this study's DQN reinforcement learning recorded more accurate temperature control results (T_{mean}=25.90°C, S.D.= 0.088°C, T_{min}=25.66°C, T_{max}=26.12°C).

Acknowledgement

The authors highly appreciate the financial support from Korea Testing Laboratory (Project No. CBS1826).

References

[1] D. L. Marino, K. Amarasinghe, and M. Manic, "Building energy load forecasting using deep neural networks," in Proceedings of the IEEE Industrial Electronics Society (IECON), 2016, pp. 7046–7051.

[2] A. Rahman, V. Srikumar, and A. D. Smith, "Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks," vol. 212, Applied Energy, 2018, pp. 372–385.

[3] A. S. Ahmad, M. Y. Hassan, M. P. Abdullah, H. A. Rahman, F. Hussin, H. Abdullah, and R. Saidur, "A review on applications of ANN and SVM for building electrical energy consumption forecasting," vol 33, Renewable and Sustainable Energy Reviews, 2014, pp. 102–109.

[4] M. Manivannan, B. Najafi, and F. Rinaldi, "Machine learning-based short-term prediction of air-conditioning load through smart meter analytics," vol. 10, Energies, 2017, pp. 1905.

[5] S. K. Jeong, C. H. Han, L. Hua, and W. K. Wibowo, "Systematic design of membership functions for fuzzy logic control of variable speed refrigeration system," vol. 142, Applied Thermal Engineering, 2018, pp. 303–310.

[6] H. Yan, Y. Xia, X. Xu, and S. Deng, "Inherent operational characteristics aided fuzzy logic controller for a variable speed direct expansion air conditioning system for simultaneous indoor air temperature and humidity control," vol. 158, Energy and Buildings, 2018, pp. 558–568.

[7] Z. Li, X. Xu, S. Deng, and D. Pan, "A novel neural network aided fuzzy logic controller for a variable speed (VS) direct expansion (DX) air conditioning (A/C) system," vol. 78, Applied Thermal Engineering, 2015, pp. 9–23.

[8] X. Xu, Z. Zhong, S. Deng, and X. Zhang, "A review on temperature and humidity control methods focusing on air-conditioning equipment and control algorithms applied in small-to-medium-sized buildings," vol. 162, Energy and Buildings, 2018, pp. 163–176.

[9] A. Kusiak, and G. Xu, "Modeling and optimization of HVAC systems using a dynamic neural network," vol. 42, Energy, 2012, pp. 241–250.

[10] P. V. Fazenda, K. Veeramachaneni, P. Lima, and U-M. O'Reilly, "Using reinforcement learning to optimize occupant comfort and energy usage in HVAC systems," vol. 6, J. of Ambient Intelligence and Smart Environments, 2014, pp. 675–690.

[11] K. Yan, C. Zhong, Z. Ji, and J. Huang, "Semi-supervised learning for early detection and diagnosis of various air handling unit faults," vol. 181, Energy and Buildings, 2018, pp. 75–83.

[12] W. Vallandares, M. Galindo, J. Gutierrez, W. C. Wu, K. K. Liao, J. C. Liao, K. C. Lu, and C. C. Wang, "Energy optimization associated with thermal comfort and indoor air control via a deep reinforcement learning algorithm," vol. 155, Building and Environment, 2019, pp. 105–117.

[13] C. C. Cheng, and D. Lee, "Artificial intelligence-assisted heating ventilation and air conditioning control and the unmet demand for sensors: Part 1. Problem formulation and the hypothesis," vol. 19, Sensors, 2019, pp. 1131.

[14] M. S. Yoon, J. Lim, T. S. Qahtani, and Y. Nam, "Experimental study on comparison of energy consumption between constant and variable speed air-conditioners in two different climates," in Proceedings of the Asian Conference on Refrigeration and Air-conditioning (ACRA), 2018, E342.

[15] J. Lim, M. S. Yoon, T. S. Qahtani, and Y. Nam, "Feasibility study on variable-speed air conditioner under hot climate based on real-scale experiment and energy simulation," vol. 12, Energies, 2019, pp. 1489.

[16] R. S. Sutton, and A. G. Barto, Reinforcement learning: An introduction (2nd edition). MIT Press, Cambridge, MA, 2018.

[17] R. Bellman, Dynamic programming. Princeton University Press, 1957.

[18] C. J. Watkins, Learning from delayed rewards, PhD thesis, King's College, Cambridge, England, 1989.

[19] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," vol. 36, Biol. Cybernetics, 1980, pp. 193–202.

[20] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," vol. 115, Int. J. of Computer Vision, 2015, pp. 211–252.

[21] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," vol. 27, Advances in NIPS, 2014, pp. 2672–2680.

[22] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back propagating errors," vol. 323, Nature, 1986, pp. 533–536.

[23] H. Sak, A. W. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," in Proceedings of Interspeech, 2014, pp. 338-342.

[24] H. Sak, A. W. Senior, K. Rao, and F. Beaufays, "Fast and accurate recurrent neural network acoustic models for speech recognition," in Proceedings of Interspeech, 2015, pp. 1468-1472.

[25] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," arXiv preprint arXiv:1312.5602, 2013.

[26] I. A. Hosu, and T. Rebedea, "Playing atari games with deep reinforcement learning and human checkpoint replay," arXiv preprint arXiv:1607.05077, 2016.

[27] ISO 5151, "Non-ducted air conditioners and heat pumps–Testing and rating for performance," International Organization for Standardization, Geneva, Switzerland, 2017.

[28] ASHRAE 116, "Methods of testing for rating seasonal efficiency of unitary air-conditioners and heat pumps," American Society of Heating, Refrigerating and Air-conditioning Engineers, Atlanta, GA, USA, 2010.

[29] M. Abadi, et. al, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, https://www.tensorflow.org

[30] A. Afram, and F. J. Sharifi, "Effects of dead-band and set-point settings of on/off controllers on the energy consumption and equipment switching frequency of a residential HVAC system," vol. 47, J. of Process Control, 2016, pp. 161–174.