Advanced Optimal Control of Indoor Environmental Devices for Indoor Air Quality Using Reinforcement Learning

Sun Ho Kim[1], Jeong Won Kim [1], and Hyeun Jun Moon[*1]

1 Dankook University
152, Jukjeon-ro, Suji-gu, Yongin-si, Gyeonggi-do, Republic of Korea

ABSTRACT

This study aims to develop and evaluate an advanced control method for acceptable indoor air quality (e.g., particulate matter and CO_2) with low energy consumption in a residential space. A ventilation system, an air purifier, and a kitchen hood system are installed in the testbed to maintain a healthy IAQ. To accomplish the objective, we use a double deep Q-network (DDQN) which is one of the reinforcement learning. This study utilizes a co-simulation platform with EnergyPlus and Python. The optimal control model was trained for 5 days to represent various outdoor conditions and indoor living contexts in residential buildings by introducing emission rates of the indoor fine particles according to occupant's activities. The evaluation of the suggested optimal control was performed by comparison with a simple on/off method for environmental devices. As a result, the DDQN control showed an improvement of 2.5% (PM 2.5) and 0.6% (CO_2) of healthy air ratio while reducing 45.5% of energy consumption.

KEYWORDS

Reinforcement learning, Integrated control scheme, Indoor air quality, Ventilation, Double deep Q-network

1 INTRODUCTION

As most people in developed nations spend more than 90% of their time indoors, indoor air quality (IAQ) has an important role in and a huge impact on protecting occupants' health, morale, working efficiency, productivity, and satisfaction. We can use a ventilation system, an air purifier, and a kitchen hood, etc. in a residential environment to improve and maintain acceptable IAQ (PM 2.5 and CO_2). Earlier studies showed control methods for environmental devices, which could maintain favourable conditions in terms of IAQ (Kim et al., 2020). However, this method used a simple on/off control to maintain the indoor concentration of fine particulates (PM 2.5) and CO_2 within a defined upper limit. This simple control method can cause inefficient building operation because it does not reflect the changes of indoor–outdoor environmental conditions, the operation status of the environmental devices, and occupants' activities. To overcome these limitations, we suggested a new advanced control method with a double deep Q-network (DDQN), which uses a data-driven approach to find the optimal control of several environmental control devices to maintain IAQ with low energy consumption.

2 INDOOR AIR QUALITY GUIDELINES

Pollutants affecting IAQ can be divided into 15 substances such as carbon dioxide, carbon monoxide, formaldehyde, radon, ozone, and particulate matter (Jones, 1999). In this study, particulate matters and CO_2 were used as IAQ indicators. The guideline for particulate matters varies according to the institutions and countries. As shown in Table 1, to maintain healthy IAQ, the US Environmental Protection Agency (EPA) suggests that PM 10 should be under

150µg/m^3 and PM 2.5 is under 35µg/m^3 on the 24-hour average (US. EPA, 2013). Also, the World Health Organization (WHO) suggests that PM 10 is under 50µg/m^3 and PM 2.5 is under 25µg/m^3 on the 24-hour average (WHO, 2005). In this study, we set a limitation of indoor particulate matter as 25 µg/m^3 for PM 2.5 to satisfy both guidelines.

Table 1: Guidelines for indoor particulate matter

Institution	PM 10	PM 2.5
US EPA	≤150µg/m^3	≤35µg/m^3
WHO	≤50µg/m^3	≤25µg/m^3

The acceptable level of indoor CO_2 concentration varies from 700 ppm to 5000 ppm depending on the country, standards, buildings, and standards (Lowther et al., 2021). As shown in Table 2, our study selects 1000 ppm as an acceptable level of indoor CO_2 concentration, because this value has no adverse effect on the occupant's health and serves as a standard for adequate ventilation of the room.

Table 2: Guidelines of indoor CO_2 concentrations

CO_2 Guideline concentration	Country	Standard	Description
1000ppm	UK	British Standard (BS EN 16798-1:2019)	Good indoor air quality (residential and non-residential)
	US	US EPA Facilities Manual Vol 2: Architecture and Engineering Guidelines	8 h average
	China	GB/T 18883-2002, Indoor air quality standard. Standards Press of China	24 h average (0.1% CO_2 = 1000 ppm)
	Korea	Korea Occupational Safety and Health Agency (KOSHA), Guideline development for evaluation and management of office air quality (II)	8 h average (office)

3 INDOOR AIR QUALITY CONTROL ALGORITHM

3.1 Rule-based control scheme

Figure 1 is a flowchart of a control scheme including indoor particulate matter and CO_2 guidelines to maintain acceptable IAQ. The control scheme uses the simple on/off method based on an upper limit of indoor PM 2.5 and CO_2 concentrations. When indoor CO_2 concentration exceeds the acceptable level (1000 ppm), the ventilation system is operated to decrease the indoor CO_2 concentration. The rule-based control scheme operates the ventilation system, air purifier, and kitchen hood to remove indoor particulate matter when indoor PM 2.5 concentration exceeds the acceptable level of 25µg/m^3. Finally, if both concentrations of CO_2 and PM 2.5 satisfy the criteria, all environmental devices are turned off.

3.2 Advanced optimal control (Double Deep Q-network)

The indoor environment is affected by various influencing factors, such as outdoor conditions, the operation status of indoor environmental devices, occupants' activities, and many others (Frontczak and Wargocki, 2011). However, simple rule-based control cannot reflect the complexity of influencing factors (Shaikh et al., 2014). To overcome this limitation, we developed an optimal control by employing Double Deep Q-network (DDQN).

DDQN is derived from Deep Q-network (DQN). DQN combines reinforcement learning with

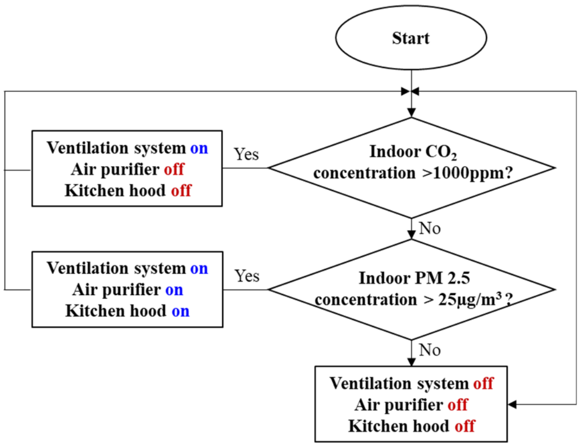

Figure 1: Flowchart of the rule-based control scheme

a class of artificial neural networks known as deep neural networks. The Q-network is updated to minimize the mean square error with maximum value from the target Q-network by using Equation (1).

$$L_I(\theta_I) = \mathbb{E}[(r + \gamma \max_{a'} Q(s', a'; \theta_i^-) - Q(s, a; \theta_i))^2] \tag{1}$$

However, DQN were found to overestimate the action value, leading to poorer policies (Van Hasselt, 2011). To overcome this limitation, van Hasselt et al. proposed the DDQN algorithm (Van Hasslet, Guez, and Silver, 2016). In DDQN, the current Q-network is used to select the next greedy action, and the target network evaluates the selected action. The loss function of DDQN can be described by Equation (2).

$$L_I(\theta_I) = \mathbb{E}[(r + \gamma Q(s', \arg\max_{a'} Q(s', a'; \theta_i^-) - Q(s, a; \theta_i))^2] \tag{2}$$

To train optimal control by employing DDQN, we selected states, actions, and rewards as shown in Table 3. In this study, 12 states were selected to describe the indoor environment, outdoor environment, occupant's activities, and the operation statuses of indoor environmental devices. We could select the control action for the ventilation system, the air purifier, and the kitchen hood. There are 4 possible actions for the ventilation system and the air purifier, and three for the kitchen hood.

Table 3: State, actions, reward for DDQN

State	Action(m^3/min)		Reward
Date(-) Time(-) Occupancy activity(-) Outdoor concentration of PM 2.5(μg/m^3) Indoor concentration of PM 2.5(μg/m^3) Emission rate of PM 2.5(μg/min) Outdoor concentration of CO_2(ppm) Indoor concentration of CO_2(ppm) Emission rate of CO_2(m^3/s) Air flow rate of ventilation system(m^3/min) Air flow rate of kitchen hood(m^3/min) Air flow rate of air purifier(m^3/min)	Ventilation system (flowrates)	0, 2.5, 3.4, 4.2	r_{EC}, Energy consumption (kWh) r_{PM}, Indoor concentration of PM 2.5(μg/m^3) r_{CO2}, Indoor concentration of CO_2 (ppm)
	Air purifier (flowrates)	0, 2.5, 3.2, 4.8	
	Kitchen hood (flowrates)	0, 2.7, 3.3	

As shown in Equation (3), three reward factors, r_{EC}, r_{PM} and r_{CO2}, are used to consider IAQ and energy consumption at the same time. Equations (4) and (5) represent the rewards for indoor concentration of PM 2.5 and CO_2. If each IAQ factor is below an acceptable level, a positive reward of 1 is provided because healthy IAQ was achieved. On the contrary, when an each IAQ factor is over the acceptable level, a reward of −1 is provided to impose a penalty. The reward for energy consumption (r_{ec}) includes the electrical energy used by the ventilation system, air purifier, and kitchen hood. This reward is provided as a penalty in r_t to minimize energy consumption.

$$r_t = r_{EC} + r_{PM} + r_{CO2} \tag{3}$$

$$r_{PM} = \begin{cases} +1\ If\ indoor\ PM\ 2.5\ concentration\ is\ below\ 25\mathrm{\mu g/m^3} \\ -1\ If\ indoor\ PM\ 2.5\ concentration\ is\ over\ 25\mathrm{\mu g/m^3} \end{cases} \tag{4}$$

$$r_{CO2} = \begin{cases} +1\ If\ indoor\ CO_2\ concentration\ is\ below\ 1000\ \mathrm{ppm} \\ -1\ If\ indoor\ CO_2\ concentration\ is\ over\ 1000\ \mathrm{ppm} \end{cases} \tag{5}$$

The timestep for the EnergyPlus simulation was set to 60 per hour or one-minute steps. This means 1440 simulations were performed on EnergyPlus per one day. In this study, EnergyPlus running for five days was regarded as one episode, and 3000 episodes were iterated to explore the optimal DDQN policy.

4 METHOES

4.1 Building Integrated Control Testbed (BICT)

In this study, a simulation model was created for the building-integrated control testbed (BICT) at Dankook University in Yongin, Korea. The BICT is an experimental chamber that consists of an air conditioner, a ventilation system, an air purifier, a kitchen hood, a humidifier, various sensors to monitor the indoor and outdoor environmental conditions, and meters to measure energy consumption as well. The exterior of the BICT and environmental control devices are shown in Figure 2. Table 4 shows the construction and configuration of the BICT, along with detailed information on the environmental control systems.

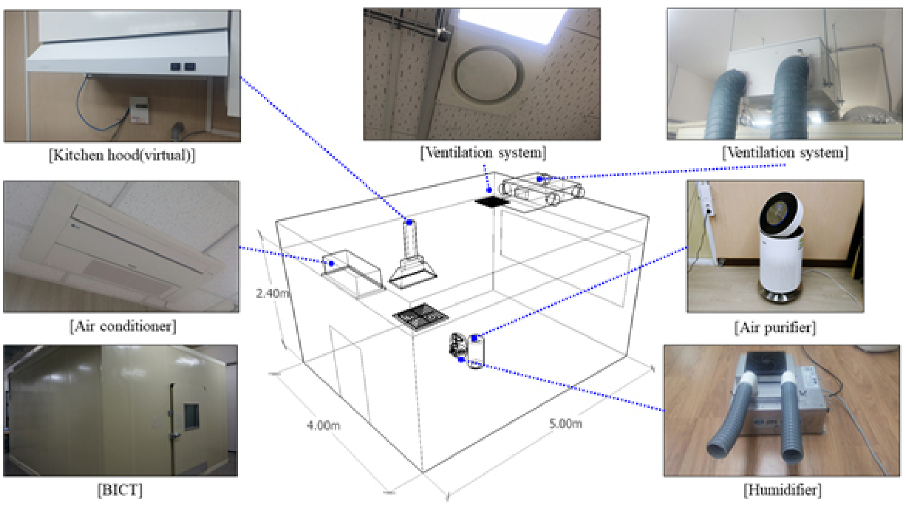

Figure 2: Floor plan of the BICT and indoor environmental control devices.

Table 4: Virtual testing system configurations

BICT Envelope	Size	4.0 m × 5.0 m × 2.4 m	
	Materials	Laminate floor on concrete and urethane layers	
		Urethane panel with gypsum lapping	
		Double-glazed window with 5 mm glass panes and 5 mm air cavity	
Environmental Control Systems	Ventilation system	Supply airflow rate(Max flow rate)	0.07 m^3/s
		Exhaust airflow rate(Max flow rate)	0.07 m^3/s
		Rated power	400 W
	Air purifier	Clean airflow rate(Max flow rate)	0.08 m^3/s
		Rated power	30W
	Kitchen hood	Exhaust airflow rate(Max flow rate)	0.06 m^3/s
		Rated power	50W

4.2 Co-simulation platform

As shown in Figure 3, the suggested control algorithms were constructed and evaluated in a co-simulation platform between the EnergyPlus and the Python. The EnergyPlus was utilized to simulate indoor CO_2 concentrations and energy consumptions according to occupancy activities and the operation statues of the ventilation system and the kitchen hood. However, there is no module to simulate the concentrations of indoor particulate matters in the EnergyPlus. Thus, the Nazaroff equation was implemented using the python language to calculate the concentrations of indoor particulate matters. The Python module 'eppy' was utilized to connect the control actions for the DDQN algorithm and the EnergyPlus building simulation program (Philip, 2019). Also, the DDQN was implemented on the library Keras. When the current state values simulated from the EnergyPlus are transferred to the Python, the DDQN factors derive the optimal control actions that satisfies IAQ with low-energy consumption based on the input state.

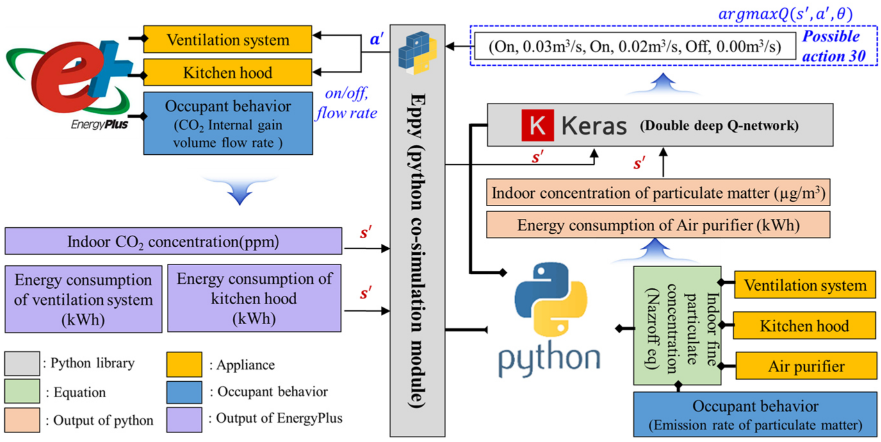

Figure 3: Co-simulation platform with EnergyPlus and Python

4.3 Dynamics of Indoor Particulate Matters

In this study, we utilized an indoor particle dynamic to calculate indoor particulate concentration as shown below (Nazaroff, 2014).

$$d(C_i V)/dt = E + C_O[Q_S(1 - \eta_s + Q_N + Q_L P)] - C_i[Q_F \eta_F + \beta V + (Q_S + Q_N + Q_L + Q_H)] \quad (6)$$

The equation (6) is to calculate indoor concentrations of fine particulate in an enclosed space. In this equation, the E(µg/min) is an emission rate of indoor fine particulates. The emission rate varies widely according to occupancy activities; thus, two typical behaviors have been selected in this study to calculate indoor fine particulate concentrations: Vacuuming, and Cooking (Oven/Grilled/Fried). V(m^3) is the volume of the room. C_o ($\mu g/m^3$) is an outdoor air concentration of the particulate matter. An outdoor particulate matter enters the room by three pathways: mechanical supply, Q_s (m^3/min); natural ventilation, Q_N (m^3/min); and infiltration, Q_L (m^3/min). η_S is a filter efficiency which located in mechanical supply path. Q_F and Q_H (m^3/min) are an additional flow path. In this study, Q_F is flow rate of an air purifier which passes through filter with efficiency η_F, and Q_H is exhaust air from the room by kitchen hood. P(-) is a fraction of particles in the infiltration flow path. Finally, β(-) is the deposition rate onto the room surfaces. Table 5 shows the selected values of each parameter in our study for calculation of fine particulate concentrations. We studied references to set the value of emission rates(E) (He et al., 2004; Hu et al., 2012), fraction of particles(P) and deposition rate(β) (Kim, 2018). Other values were acquired directly from the BICT.

Table 5: Input value for concentration of fine particulate (PM 2.5)

E(µg/min)			V (m^3)	Co (µg/m^3)	Q_S (m^3/min)	η_S (-)	Q_F (m^3/min)	η_F (-)	Q_H (m^3/min)	Q_N (m^3/min)	Q_L (m^3/min)	P (-)	β (min^{-1})
Vacuuming		70	68	6µg/m^3 ~ 235 µg/m^3	4.2	0.9	4.8	0.9	2.4	0	0.56	0.7	0.0067
Cooking	Oven	10											
	Grilled	283											
	Fried	1483											
Others		0											

4.4 Occupant’s activities

In this study, the occupant’s activities were divided into seven categories (sleep, resting, cooking (oven, grilled, fried), eating, vacuuming, working, exercising) considering the emission rates of particulate matters and CO_2 concentrations for each activity. Table 6 shows the emission rates of PM 2.5 and CO_2 concentrations based on the occupant’s activity (He et al., 2004; Hu et al., 2012; U.S. DOE., 2019). The occurrence and duration time of each occupant’s activity was studied from literature. This study set the occupancy schedules from the ICATUS 2016 report (United Nations Statics Division, 2017) and the Time use survey 2019 (Statics Korea, 2020). Figure 4 shows the emission rate of PM 2.5 and CO_2 concentrations based on the occupant’s activities.

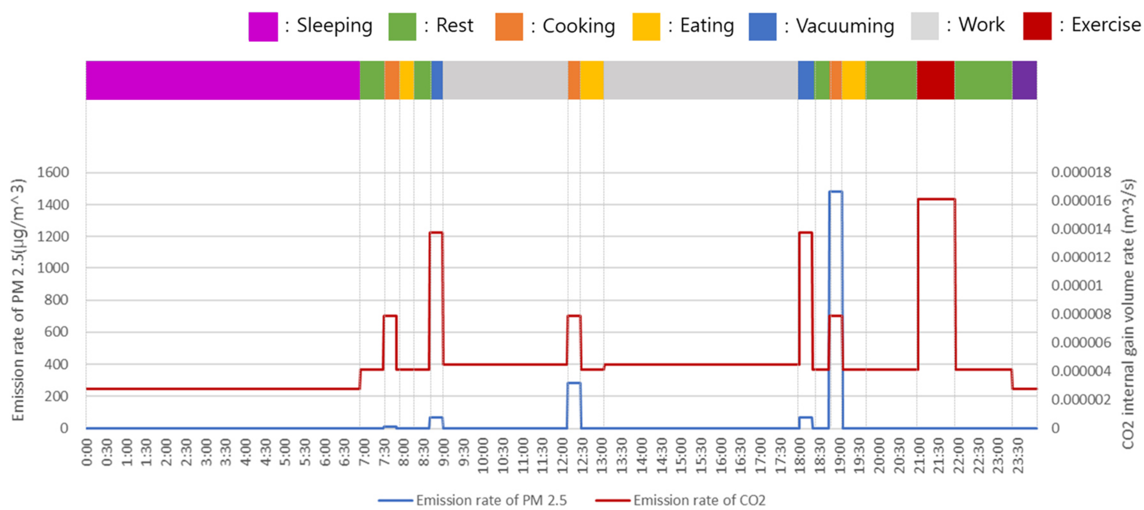

Figure 4: Occurrence and duration time of occupancy activities

Table 6: Emission rate of PM 2.5 and CO_2 concentrations according to occupant's activities

		Sleeping	Exercise	Vacuuming	Cooking			Eating	Rest	Work
					Oven	Grilled	Fried			
PM 2.5	Emission rate of PM 2.5 (μg/min)	0	0	70	10	283	1483	0	0	0
CO_2	Number of people (-)	1								
	Activity level (W)	72	423	360	207			108	108	117
	Emission rate of CO_2 (m^3/s)	1.65e-4	9.69e-4	8.25e-4	4.07e-4			2.47 e-4	1.86 e-4	2.68 e-4

5 RESULTS

The results from the suggested advanced optimal control using the DDQN algorithm could be compared to the rule-based approach in terms of energy consumption (kWh), and healthy air ratio (%) of PM 2.5 and CO_2. Total energy consumption is the sum of energy consumption of the ventilation system, the air purifier, and the kitchen hood. Like as equation (7), the healthy air ratio is defined as the ratio of the time duration under the acceptable level of PM 2.5 and CO_2 to the reference time duration (5 days).

$$\text{Healthy air ratio (\%)} = \frac{\text{Duration under acceptable level (PM 2.5, CO}_2\text{) in minutes}}{(5\times24\times60)\text{ minute}} \quad (7)$$

Table 7: Comparison of Rule-based control and DDQN

Evaluation factor		**Rule-based control**	**DDQN***	**Improvement**
Energy consumption (kWh)	Ventilation system	1.15	0.15(±0.06)	-1.00
	Air purifier	0.25	0.50(±0.02)	+0.25
	Kitchen hood	0.33	0.37(±0.08)	+0.04
	Total	1.73	1.02(±0.13)	-0.71
Healthy air ratio (%)	PM 2.5	93.1	95.6(±0.09)	+2.5
	CO_2	99.3	99.9(±0.08)	+0.6

*Performance of DDQN expressed as averaged value(±std) of last 50 episodes.

Table 7 expresses the performance of the suggested advanced optimal control as the average value of the last 50 episodes in DDQN learning. In terms of energy consumption, the total energy consumption of the suggested control was 1.02 kWh, which is 45.5% lower than the energy consumption from the rule-based control scheme (1.73 kWh). More specifically, in the case of DDQN, the air purifier and the kitchen hood consumed slightly more energy (the air purifier: 0.25 kWh, the kitchen hood: 0.04 kWh) than the rule-based control scheme. However, this increase was offset by the decreased energy consumption of the ventilation system. In the operation of the ventilation system, the optimal control consumed 64.1% less energy than the rule-based control scheme. This operation showed the advanced optimal control reflected the ventilation system's characteristics that consumes relatively high energy compared to the air purifier and kitchen hood. In other words, as shown (a) in Figure 5, the advanced optimal control only operated the ventilation system when occupancy activity with high emission rates of PM 2.5 and CO_2 concentrations such as cooking and exercising. The decrease in removal of PM 2.5 due to a reduction in operating time of the ventilation system was offset by increasing the operation of the air purifier and the kitchen hood as shown in (b) and (c) in Figure 5. This shows that the DDQN control learned the availability of operating the indoor environmental devices according to the indoor and outdoor environment and occupancy activity to reduce total energy consumption.

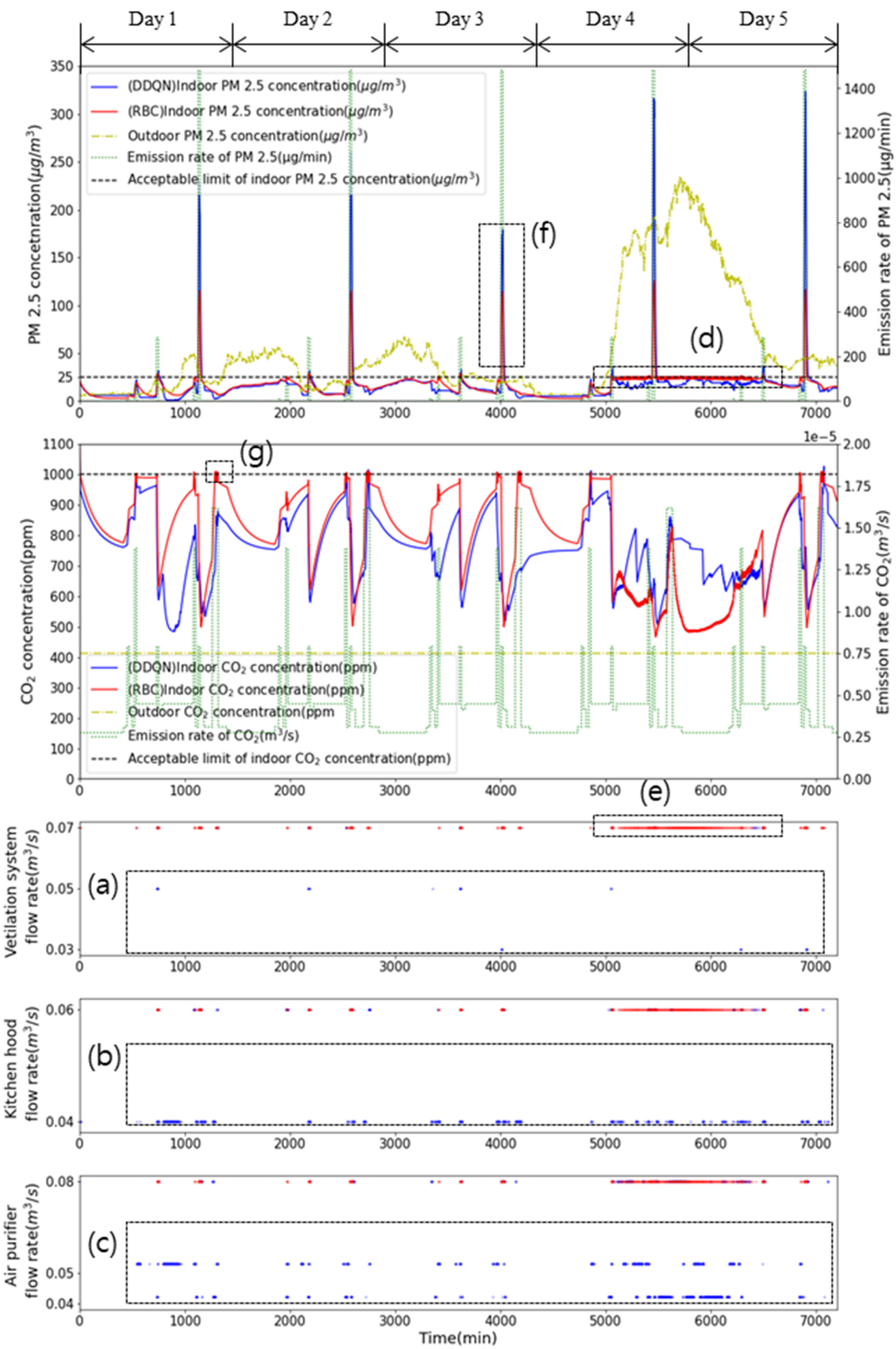

Figure 5: Concentrations of PM2.5, CO_2 and environmental device operation status

The healthy air ratio (PM 2.5) of the advanced optimal control was 95.6%, which was 2.5% higher than the rule-based control scheme. In the rule-based control, the indoor concentration of PM 2.5 was maintained above the acceptable level because the ventilation system, air purifier, and kitchen hood were operated after exceeding the acceptable level, 25μg/m^3. This inefficient operation of the rule-based control scheme could be shown on day 4 to day 5 like as (d) and (e) in Figure 5, which is outdoor particulate matter concentration was very unhealthy. On the contrary, optimal control operated environmental devices before the indoor concentration of PM 2.5 exceeds the upper limit of the acceptable level as shown in the figure, and this efficient operation led to an increase in the healthy air ratio of PM 2.5. However, both the DDQN and the rule-based control scheme were not able to maintain the acceptable level of PM 2.5 when the occupant generated a large amount of particulate matters such as cooking (Fried) as shown in (f) in Figure 5. This result provide evidence this residential space needs additional measures (e.g., opening windows, installation of high-efficiency filters, etc.) to maintain appropriate IAQ.
The healthy air ratio (CO_2) of the DDQN control was 99.9%, which was 0.6% higher compared to control scheme. As shown in Figure 5, the indoor CO_2 concentrations exceeded the acceptable level (1000 ppm) when the occupant's activity level was high, such as cleaning and exercising as shown in (g) in Figure 5. The DDQN control showed an improved healthy air ratio compared to the rule-based control scheme by operating the ventilation system before the indoor concentration of CO_2 exceeds the upper limit of the acceptable CO_2 level.

6 CONCLUSION

In this study, we proposed an advanced optimal control algorithm based on reinforcement learning to maintain healthy IAQ with low energy consumption. In terms of energy consumption, the DDQN control showed a 45.5% reduction compared to the rule-based control scheme. This is because the DDQN control actively uses the air purifier and the kitchen hood that consume relatively lower energy compared to the ventilation system. In terms of healthy air ratio, the DDQN control showed improved performance by 2.5% at PM 2.5. This study showed that the advanced control with reinforcement learning could reflected the indoor-outdoor environmental conditions, the operation status of the environmental control devices, and occupant's activities. The suggested approach could be used to maintain acceptable IAQ while reducing energy consumptions in residential buildings.

7 ACKNOWLEDGEMENTS

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea (No. 20212020800120).
This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea (No. 20212020800050).

8 REFERENCES

C. He, L. Morawska, J. Hitchins, D. Gillbert. (2004). *Contribution from indoor sources to particle* number and mass concentrations in residential houses, Atmospheric environment 38, pp. 3405-3415.

M. Frontczak., P. Wargocki. (2011). Literature survey on how different factors influence human comfort in indoor environments. Build. Environ, 46, 922–937.

J. H. Kim. (2018). Development of prediction model for indoor fine particulate, Magazine of the SAREK 47(12), pp. 44-48.
A.P. Jones. (1999). Indoor air quality and health. Atmospheric Environment, 33(28), 4535-4564.
S. D. Lowther., S. Dimitroulopoulou., K. Foxall., C. Shrubsole., E. Cheek., B. Gadeberg., O. Sepai. (2021). Low-Level Carbon Dioxide Indoors—A Pollution Indicator or a Pollutant? A Health-Based Perspective. Environments, 8(11), 125.
S. Philip. (2019). Eppy Documentation; Github Repository, Available online: https://pypi.org/project/eppy.
P.H. Shaikh., N.B.M. Nor., P. Nallagownden., I. Elamvazuthi., T. Ibrahim. (2014). A review on optimized control systems for building energy and comfort management of smart sustainable buildings. Renew. Sustain. Energy Rev. 2014, 34, 409–429.
Statistics Korea, (2020). 2019 Time Use Survey.
S.H. Kim., Y.R. Yoon., J.W. Kim., H.J. Moon. (2020) An integrated operation of indoor environmental systems considering occupants behaviour to maintain thermal comfort and the concentration of indoor particulate matters, The Korean Society of Living Environmental System
T. Hu, B.C. Singer, J.M. Logue. (2012). Compilation of Published PM2.5 Emission Rates for Cooking, Candles and Incense for Use in Modeling of Exposures in Residences. LBNL report.
United States, Department of Energy. (2019). EnergyPlus Input output reference.
United Nations Statistics Division. (2017). International Classification of Activities for Time Use Statistics 2016 (ICATUS 2016).
United States, Environment Protection Agency (2013). NAAQS Table from Website:https://www.epa.gov/criteria-air-pollutants/naaqs-table.
H. Van Hasselt. Insights in Reinforcement Learning: Formal Analysis and Empirical Evaluation of Temporal-Difference Learning Algorithms. Ph.D. Thesis, Utrecht University, Utrecht, The Netherlands, 2011; ISBN 9789039354964.
H. Van Hasselt., A. Guez., D. Silver. (2016). Deep reinforcement learning with double Q-Learning. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16).
Word Health Organization, (2005). WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide.
W.W Nazaroff. (2014). Indoor particle dynamics, Indoor air 2014, pp. 175–183.