Control Strategies and Thermal Performance of Responsive Building Envelope with Active Insulation System

Yawen He
Student Member ASHRAE

Farbod Farhimi, PhD

Hongyu Zhou, PhD

ABSTRACT

Thermally dynamic building envelope is a promising technology to achieve building energy saving while improving thermal comfort. Their performance is highly dependent on the local climate conditions as well as on the way the dynamic properties are operated/controlled. The evaluation of whole building performance through building energy simulation can be useful to understand the potentials of different dynamic opaque envelope with active insulations in a specific context. This paper evaluates the potential to use model-free online reinforcement-learning (MFORL) control to regulate the behavior of dynamic building envelopes. Specifically, two control strategies were formulated and evaluated on dynamic opaque envelopes that consist of a concrete layer sandwiched between two active insulations on both sides of the thermal mass: (i) simple temperature-driven rule-based control, and (ii) MFORL control. The controllers were preliminarily tested in two scenarios with 10-day representative behavior under mild climate or during transitional seasons. The results show that MFORL control is promising in achieving adaptive thermal behavior for dynamic building envelopes and may have advantages over traditional rule-based controllers under complex environment.

INTRODUCTION

Responsive Building Envelopes

The dynamic nature of forces and energies acting on our building stokes has prompted the emergence of high-performance building skins that are interactive and responsive to the environment (Wigginton and Harris, 2002; Stec and Paassen, 2005; Biloria and Sumini, 2007; Joe et al., 2013; Loonen et al., 2013; Loonen, 2015). Increasingly, building façades are being developed as complex systems of material assemblies attuned to climate and energy optimization – they are equipped with new performative materials, sensors, actuators, and artificial intelligence that support automated and dynamic functionalities of buildings, such as regulating natural day lighting, air and sound transmission, thermal transfer, and interior air quality (Velikov K, 2013). This paradigm shift from static building envelope to more intelligent 'building skin' that can sense and respond to environmental changes provides opportunities for energy saving, improving occupant comfort, and enabling adaptations to the changing climate.

To date, several approaches have been taken to achieve the "responsive building envelope (RBE)" concept including the utilization of intrinsic material properties such as the thermoresponse of bimetals and shape-memory polymers for self-

Yawen He is a PhD Candidate at the University of Tennessee, Knoxville, TN and **Hongyu Zhou** is an Associate Professor in Civil and Environmental Engineering, University of Tennessee, Knoxville, TN. **Farbod Famihi** is an Associate Professor in Mechanical and Aerospace Engineering at the University of Alabama in Huntsville, Huntsville, AL.

ventilation and day-lighting control (Sung, 2010; Brigham, 2015; Rybkowski et al., 2015). In addition, recent research and development efforts have advanced materials and devices to achieve switchable or variable thermal properties in building envelope assemblies. Inspired by the response of animal skins to thermal environment variations, these variable insulations are designed to selectively transfer heat across the envelope, making it possible to insulate heat flux and dissipate/absorb heat on demand (Cui and Overend, 2019). Examples of "on-and-off" thermal switch include changing material thermal conductivity through hydration/dehydration, which leads to micro-structure change (Tomko *et al.*, 2018); changing porosity; materials orientation (Wu *et al.*, 2014) that can respond in a few seconds. Continuously variable thermal insulation can be achieved by variable-pressure vacuum insulated panel (VIP) (Berge *et al.*, 2015). The variable thermal insulation may also be achieved through multi-layer retractable thermal insulation. Antretter et al. (Antretter, 2019) and Mumme et al. (Mumme and James, 2020) assessed the energy saving potential of several configurations of RBE with controllable active insulation systems (AISs), and it was found that significant reduction of total building energy consumption for heating and cooling was achieved in all climate zones.

RBE Control Strategies

The control of RBEs can be generally categorized into (i) intrinsic (passive) and (ii) extrinsic (active or semi-active) control schemes (Loonen *et al.*, 2013). Passive control strategies utilize the intrinsic properties of materials or mechanisms that are automatically triggered by a stimulus (surface temperature, solar radiation, etc.). This level of intelligence is embedded in the material and the switching mechanism is activated by a variation in its internal energy. The intrinsic control is also referred to as "direct" or "open-loop" control. In contrast, extrinsic control refers to the presence of an external decision-making component that is able to trigger the adaptive mechanisms according to a feedback rule. This is so-referred as to the feedback (or closed-loop) control type. Common control strategies included: (i) rule-based control, (ii) model predictive control (MPC) and (iii) model-free control. For RBE applications, a number of studies have explored the rule-based control strategy for RBEs as demonstrated in (Shekar and Krarti, 2017; Rupp and Krarti, 2019). While model-free control approaches have been mentioned in the scoping study for building control conducted by Yoon *et al* and Pinto *et al* (Yoon and Moon, 2019; Pinto *et al.*, 2021), their implementation and research within RBE systems have been very limited.

Reinforcement learning (RL) control is an application of a branch of machine learning that obtains an optimal control laws by maximizing a numerical delayed reward. RL control obtains the optimal control law without supervised learning and only by evaluating the immediate feedback from environment (i.e. the control system). The problem to be solved by reinforcement learning is normally defined as a Markov Decision Process (MDP), which is generally represented as a tuple (S, A, P, R), i.e. state, action, transition probability and reward function. Together with these elements, control system/environment and controller/agent comprise the basics of the whole problem. There are mainly two categories of RL algorithms for the agent, namely model-based RLs and model-free RLs. Model-based RLs find the optimal policy based on the model developed for the environment. However, learning characteristics of the whole environment and an accurate model might not necessarily provide better performance under dynamic conditions in real-time application. Model-free control refers to adaptive controllers which do not rely on any mathematical model of the real system (Michailidis *et al.*, 2018). Model-free RLs are often more flexible in updating agents under fluctuated conditions. When the external environments are unknown to the agent, online control attempts to find optimal policy/control law directly and eliminates the need for detailed modeling of the environment. These controllers generate control laws for the system's control inputs solely based on online state measurements collected from the system.

For RBEs, since some responsive building envelopes achieve their dynamic thermal properties through complex mechanisms that are difficult to be represented by explicit physics-based models. The derivation of such physic-based models is time consuming and are computationally expensive. Model-free approaches provide possibilities to eliminate the need for formulating physics-based models (Liu and Henze, 2006). Some model-free control approaches such as reinforcement learning are based on neural network (NN (Zheng and Wang, 2002)) algorithms that support online learning, so that plug-and-play (Baldi et al., 2015; Michailidis et al., 2018) controllers may be developed without computationally expensive offline training.

This paper discusses different control strategies for responsive building envelope systems. Specifically, a thermal network model is first formulated for RBEs that are comprised of a concrete thermal mass sandwiched between two layers of controllable active insulations. Two control strategies are formulated for the control of active insulations: (1) a rule-based control approach is derived simply using surface temperatures as control variables to regulate the behavior of the active insulation layers. In addition, (2) a model-free reinforcement learning (MFRL) controller is formulated for the RBE. The thermal behavior of the RBEs under these two difference control strategies are compared under different climate conditions (Miami FL and Albuquerque NM).

SIMULATION MODEL

Thermal Network Model for RBEs

A thermal network model based on a finite difference approach was developed for building envelope with dynamic thermal properties, where the discrete form of the 1-D heat transfer equation following the finite difference scheme can be expressed as:

$$\left(C_i^j + \frac{1}{2}\Delta C_i^j\right)\left(T_i^{j+1} - T_i^j\right) = \left[H_{i-1,i}^j\, T_{i-1}^j + H_{i+1,i}^j\, T_{i+1}^j - \left(H_{i-1,i}^j + H_{i+1,i}^j\right)T_i^j + Q_i^j\right]\Delta t^j \quad (1)$$

where C_i^j is the thermal capacitance of node i at j^{th} time step; $H_{i-1,i}^j$ is the heat transfer coefficient representing conduction between node i-1 and node i with variable thermal properties at j^{th} time step. For an active insulation system having negligible thermal mass (Benson, Potter and Tracy, 1994; Varga, Oliveira and Afonso, 2002; Kimber, Clark and Schaefer, 2014; Loonen, Hoes and Hensen, 2014; Wu *et al.*, 2014; Park, Srubar and Krarti, 2015; Pflug *et al.*, 2015, 2018; Tomko *et al.*, 2018), it may be represented by a 'no-mass layer' as:

$$\left(H_{i-1,i}^j + H_{i+1,i}^j\right)T_i^j = H_{i+1,i}^j\, T_{i+1}^j + H_{i-1,i}^j\, T_{i-1}^j + Q_i^j \quad (2)$$

The time-varying the heat transfer coefficient $H_{k,i}^j$ between node k and node i with variable insulation at time t^j can be controlled to regulate the amount of heat flow going through the envelope. Based on the different actuation mechanisms (and material technologies), the heat transfer coefficient may be modeled either as a binary function or a continuous variable:

$$H_{k,i}^j = f_{binary}\left(t^j, T^j, Q^j\right) = \begin{cases} H_{\max} & \text{criterion 1} \\ H_{\min} & \text{criterion 2} \end{cases} \quad \text{(3-a)}$$

$$H_{k,i}^j = f_{continous}\left(t^j, T^j, Q^j\right) \quad \text{(3-b)}$$

For this research, a dynamic building envelope layout is selected where a high thermal mass concrete layer is sandwiched between two AIS layers, see Figure 1, to allow for load shifting capabilities. In this configuration, the thermal conductivities of the active insulation layers (i.e., both exterior and interior AIS) can be separately controlled: during the heating season, the thermal conductivity of the external AIS layer can be increased when the exterior surface temperature is higher than the temperature of the thermal mass, otherwise the thermal conductivity of exterior AIS can be decreased to reduce the heat loss from the thermal mass.

Detailed thermal network model formulations for the building envelope with thermal static properties, interior partitions, and indoor air can be found in the previous work of the authors (He *et al.*, 2020). In numerical computation, long-wave radiation is linearized using the same method as detailed in (Deardorff, 1978). The whole building model was coded using Matlab Simulink. Comparisons between the simulation results of the thermal network model and EnergyPlus for surface temperatures and indoor air temperature during representative days are presented in Figure 1 (d). The simulation results of the thermal network model match closely with those obtained from EnergyPlus. The CV-RMSE (coefficient of variation of the root mean square error) and NMBE (normalized mean bias error) of all the temperature items (surface temperature and indoor air temperature) and the energy consumption (district heating or cooling) are well under the ASHRAE Guideline 14 (Monetti *et al.*, 2015) (i.e., <30% CV-RMSE and <10% NMBE for hourly data), indicating acceptable accuracy of the thermal network model.

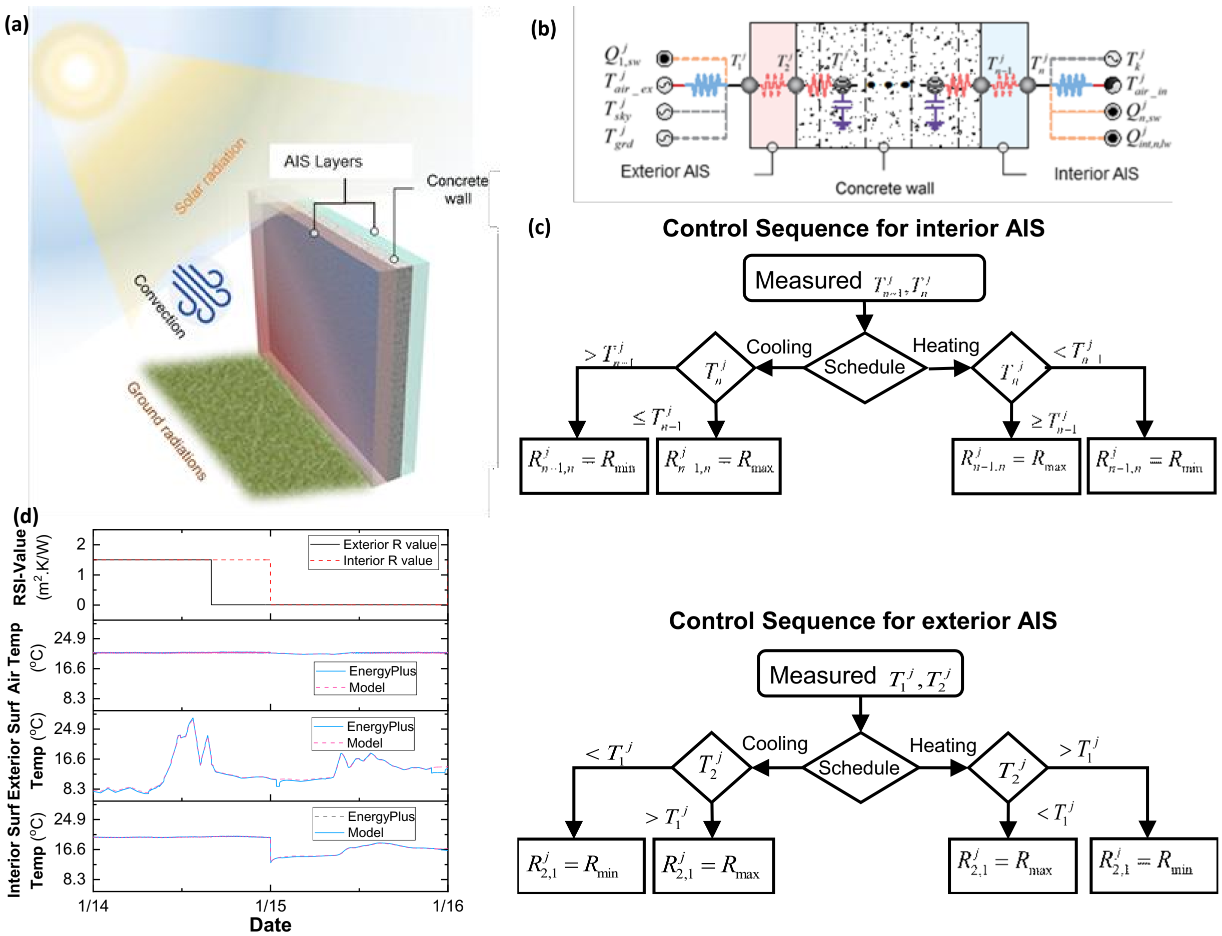

Figure 1 Configuration and control rule for RBE with concrete thermal mass sandwiched between two AIS layers: (a) illustrative showing the RBE configuration; (b) the thermal network model; (c) control rule for AIS (d) Comparison of simulation results between thermal network model and EnergyPlus

Rule-based Control

First a simple 'rule-based' controlled is formulated for the RBE as shown in Figure 1 (a). For a given moment when heating is needed and the temperature of the thermal mass is higher than the that of the interior surface, the thermal conductivity of interior AIS can be increased to enhance the heat flow. The same methodology applies to the cooling season with discharging of the thermal mass when external surface temperatures are lower than the thermal mass temperature and discharge of the thermal mass towards the zone whenever cooling is required, and the thermal mass temperature is lower than the internal surface temperature (Antretter and Boudreaux, 2019; Mumme and James, 2020). The AIS is assumed to able to change its-value between two values ('on-and-off mode'). The representative nodal diagram for the thermal network model is represented in Figure 1 (b) and an example of the control rule (for the interior AIS) is shown in Figure 1 (c).

Model-Free Online Reinforcement Learning Control

In addition, a model-free online reinforcement learning (MFORL) controller is formulated for the RBE. In general, three approaches are available for model-free reinforcement learning: policy gradient, value-based, and actor-critic. Since the actor-critic method most closely relates to the optimal adaptive control, the actor-critic method was adopted for the studies of online control for the responsive building envelope in this paper. The actor and critic are neural networks that learn the optimal control law. The actor learns to provide the best action given the current state with the policy function method. The critic learns to estimate the value of the state and the action generated by the actor. As Figure 2 shows, the actor chooses actions (thermal properties of the responsive envelope) and applies them to the environment/control system. With updated thermal properties, schedules and weather data, the environment/control system generates updated state and calculates the reward for the critic. Meanwhile, the critic identifies that action with predicted value for the current state and action pair. The critic uses the reward from the environment to determine the prediction accuracy. The critic uses the error to update itself for a better prediction. The actor also updates itself through feedback from the critic and learns to generate the correct actions.

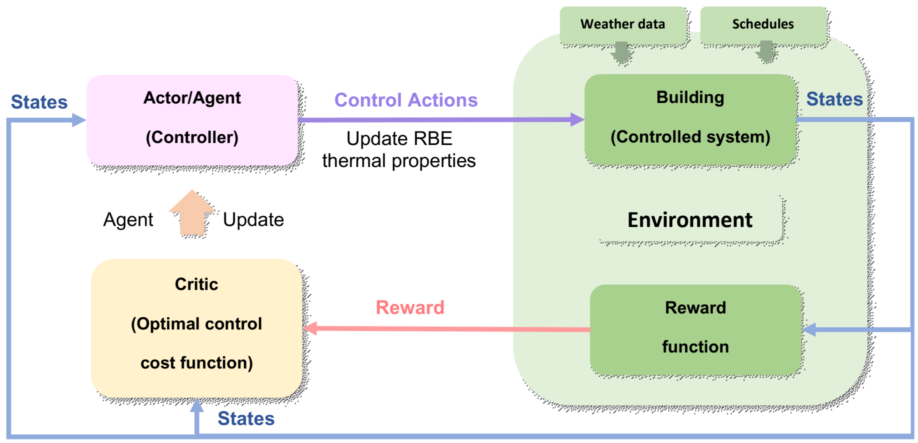

Figure 2 Schematic for reinforced learning framework for RBE

To formulate the MFORL controller, the action space contains all the possible control actions that can be taken by the agent. The state space is a set of variables related to the environment which enables the agent to learn the optimal control policy to achieve the maximum reward. In this study, the agent is expected to generate an optimal control sequence to change the thermal properties of responsive building envelope in real-time to reduce the energy consumption and improve the thermal comfort in the building at the same time. Since many parameters (e.g. the structure of the NN chosen for the actor and/or critic) affect the achievement of a successful agent, different parameter settings were applied to train the agent for optimal control using offline training. The offline training showed that the key parameters required for training a robust agent were selected correctly. With the prerequisites and the selected parameter settings by the offline training analysis, online control strategy was designed. The agent was deployed in different climate zones to test its performance and study the behavior of online training and control. The thermal performance improvement for the building envelope with static thermal properties by the RBE under rule-based control was compared with the one under model free RL control by the online trained agent.

In this study, there are two action variables: the thermal insulation of exterior AIS and interior AIS. Given the impact by dynamic characteristics of RBE and the outdoor and indoor environments, the solar radiation, outdoor and indoor air and surface temperatures, and the indoor relative humidity were chosen as the state variables to determine the optimal control actions. The simulation environment is developed using Simulink model based on the aforementioned thermal network model. The model provides the indices (e.g. surface and indoor air temperatures, PPD and AC load) for training and evaluation of reinforcement learning agent through energy simulations. Typical meteorological year 3 (TMY3) weather data were adopted for the training data for the offline training.

The objective of the agent is to minimize the energy consumption while maintaining the thermal comfort performance of the indoor space within a desired range through taking real-time actions. The reward function was designed to ensure high sensitivity to the change in the states of the system. The initial design for the reward function contained two components concerning energy consumption and thermal comfort: the AC energy consumption demand and the Predicted Percentage of Dissatisfied. Different reward functions of these two indices (square function and integral of square function) were defined and calculated to see which one is more sensitive to change in system states. The reward consisting of the square of AC energy consumption demand and the square of PPD exhibited good sensitivity during training, but the sometimes actions showed fast switching during the control. Therefore, a third component was added to the reward function to decrease the standard variation of adjacent actions. The reward function for the j^{th} step is refined though sensitivity analysis and written as:

$$\mathrm{R}^{j}=-\left\{\gamma_{1}\left(Q_{AC}^{j}\right)^{2}+\gamma_{2}\left(PPD^{j}\right)^{2}+\gamma_{3}\left[\operatorname{var}_{5}\left(R_{2,1}^{j}\right)+\operatorname{var}_{5}\left(R_{n-1,n}^{j}\right)\right]\right\} \tag{4}$$

where $\operatorname{var}_5\left(R_{2,1}^{j}\right)$ and $\operatorname{var}_5\left(R_{n-1,n}^{j}\right)$ are variance of exterior AIS and interior AIS thermal resistance within most recent 5 control sampling steps. γ_1, γ_2 and γ_3 are weights for the three components of the reward function.

Trust Region Policy Optimization (TRPO) algorithm (Schulman *et al.*, 2015, 2016; Mnih *et al.*, 2016) was used for the training of the actor-critic controller. As Table 1 shows, artificial neural networks are designed to represent the actor and critic. The node number of the input layer is set as the size of the state space and the number of the output layer for the actor is set as the size of the action space.

Table 1. Architecture of artificial neural networks

Layer	Node number (Critic)	Node number (Actor)
Input layer	20	20
Hidden layer 1	320	320
Hidden layer 2	320	320
Hidden layer 3	160	160
Hidden layer 4	160	N.A.
Output layer	1	100

The online training and control strategy with pre-training is designed for the agent deployment. For the i^{th} execution horizon of one day, two alternative agents for implementation available: one is obtained from online training using the past 1-day weather data and system behavior, and the other is a pretrained agent from the offline training. The agent with higher accumulated reward was selected as the one for execution and the weights and bias for the neural network of the agent were saved as the initial condition for online training of the next execution horizon.

Case Study

To study the thermal behavior and energy saving potential of the RBE, energy simulation is conducted for a single 13 m / 42.7 ft (L) × 8m / 24.2 ft (W) ×3.05 m / 10.0 ft (H) exterior thermal zone extracted from an intermediate floor for an apartment building. All interior walls which are assumed to be adiabatic (Jin, Favoino and Overend, 2017). The window-to-wall ratio is 15% for the exterior wall. Some other simulation parameters considered include occupant activities, lighting, equipment schedules, and infiltration level as listed in Figure 3. The RBE composed of concrete (thermal mass) layer sandwiched between two opaque AISs is used for the case study. The thermal resistance of exterior or interior AIS ranges between $R_{\min}$ equaling RSI-0.01 (R-0.06) and $R_{\max}$ equaling RSI-5 (R-28.4). To provide comparisons of the thermal behavior and energy performance of RBE as compared to the static baseline, a baseline exterior wall is designed with the same layout as RBE where a concrete mass layer is sandwiched between two rigid foam insulation panels with the same overall thermal mass as the RBE cases. The overall thermal resistance of the static baseline set the same as RBE. Typical meteorological year 3 (TMY3) weather data of Miami FL and Albuquerque NM were adopted for online training and control deployment. To study the online training and control performance of the agent, the thermal performance improvement by the RBE under model free online RL control was also compared with the one under rule-based control. For the simulation, it is assumed that the long-wave absorptivity and long-wave emissivity are the same for all wall surfaces. The indoor air temperature was controlled by HVAC with dual setpoint of 21.1℃ / 70.0 °F -23.9℃ / 75.0 °F. The Predicted Mean Vote (PMV) and Predicted Percentage of Dissatisfied (PPD) based on Frager's model (ISO, 2005) are used to evaluate thermal comfort-time performance of the building. The Long-term Percentage of Dissatisfied (LPD) is used to assess occupants' long-term thermal comfort (Carlucci, 2013).

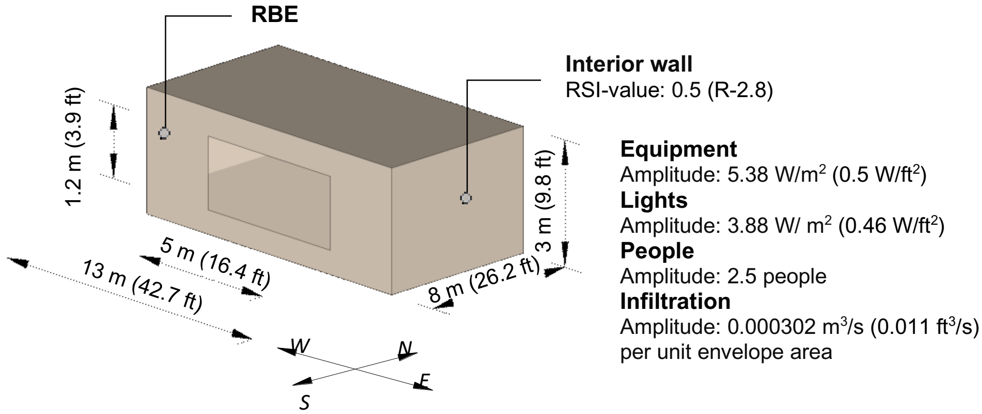

Figure 3 Simulation details

RESULTS AND DISCUSSIONS

First, the energy use for the baseline residential building is estimated for Albuquerque NM and Miami FL to assess the RBE's contribution to heating and cooling loads reduction in different climate zones during representative days. Then, the operation of the RBE with concrete thermal mass and AISs is tested for specific days to verify the implementation of both the rule-based and MFORL controllers as introduced in the previous section. In particular, weather conditions of May 9th to May 18th are selected as representative data for Albuquerque NM with high daily fluctuation in temperature, weather conditions of 16th to February 25th are selected for Miami FL to investigate how RBE makes use of benefits from outdoor environment for passive cooling.

The thermal network model is used to perform annual energy analysis when the baseline residential building is located in Albuquerque NM and Miami FL. In this analysis, static wall and roof insulations are considered with the layout same as the RBE – i.e., a concrete thermal mass layer is sandwiched between two layers of insulation foam panels with RSI-5 (R-28.4). The energy consumption of the residential building for heating and cooling is found to be 409.3 MJ / 113.7 kWh and 480.2 MJ / 133.4 kWh for Albuquerque NM and Miami FL during the representative days.

To verify the operation of both the rule-based and MFORL control strategies for the RBE when applied to residential buildings, daily temperatures as well as heating and cooling thermal load profiles are presented for the representative days: when the RSI-0.01(R-0.06) /RSI-5 (R-28.4) switching option is considered for the AIS layers. The simulation results of RBEs with rule-based control and MFORL control are presented in Figure 4 and Figure 5 for Albuquerque NM and Miami FL, respectively. Specifically, Figure 4 (a) and Figure 5 (a) illustrate the hourly variation of exterior surface temperature, exterior concrete surface temperature, interior surface temperature and indoor air temperature as well as the thermal resistance settings for both the exterior and interior AIS layers under the rule-based control. As indicated in the profiles of Figure 4 (a), During the daytime of May 9th and May 13th to May 19th, the thermal resistance of the exterior AIS layer switches to R_{max} only when the exterior surface temperature was higher than the concrete layer temperature to allow the concrete thermal mass to provide passive cooling to indoor space, as specified by the ruleset of Figure 2.

Moreover, Figure 4 (a) also compares the cooling thermal load for the rule-based controlled RBE against that obtained using the baseline static RSI-10 (R-56.8) insulation. As clearly shown in Figure 4 (a), by switching to low thermal resistance during selective hours, the AIS layer enables the charge and discharge of thermal energy into and out of the indoor space to offset a significant portion of the AC load and hence allows the building to benefit from free outdoor heating/ cooling when desired. Figure 4 (b) presents the simulation results – i.e., temperature profiles, thermal resistance settings, and sensible AC load, under MFORL control. The simulation results show that RBE under MFORL control had substantially higher AC load reduction. The energy consumption under MFORL control is nearly zero (1.1 MJ / 0.3 kWh) during these representative days as compared to that obtained under rule-based control (40.0 MJ / 11.1 kWh). Since the rule-based controller is triggered by the surface temperatures of concrete thermal mass and AIS, it worked without considering overcooling problem. Under this logic, The AIS still kept as R_{min} even when exterior surface temperature was lower than exterior concrete surface temperature when heating is needed, which undermined the beneficial thermal energy stored in the concrete layer during transitional season. Since MFORL controller does not rely on a specific ‘rule’, rather it adjusts the behavior of AIS based on the overall energy performance of the thermal zone. As a result, with longer duration of $R_{\max}$ for AISs than that under rule-based control, the concrete layer under MFORL control provided more beneficial heat flow to offset the AC load. It is worth noting that the AC demand under MFORL remained near-zero during these representative days.

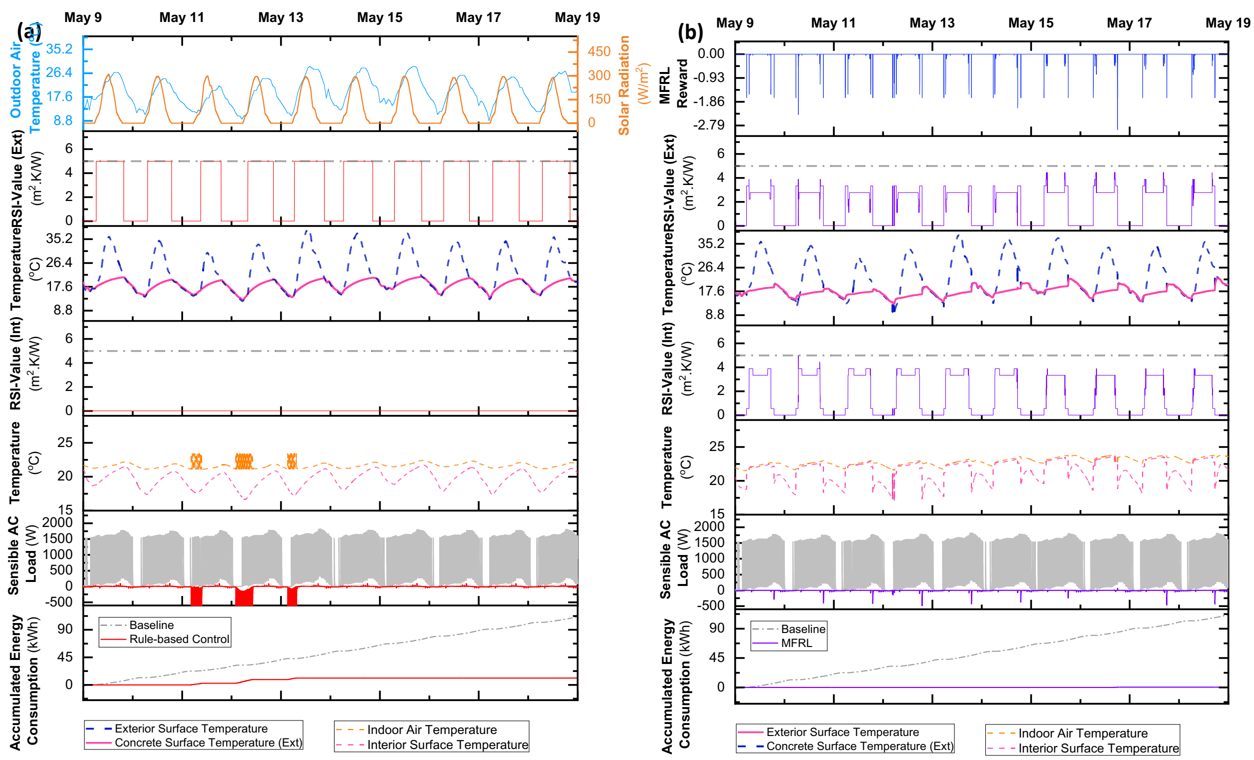

Figure 4 Simulation results for Albuquerque NM with responsive building envelope (RBE): (a) rule-based control; (b) MFORL control

Figure 5 present the simulation results obtained for Miami FL. The energy consumption under MFORL control and rule-based control are 4.7 MJ / 1.3 kWh and 158.8 MJ / 44.1 kWh, respectively. Similar to the thermal performance of RBE in Albuquerque NM, MFORL controller also provided higher AC load reduction and energy savings compared with rule-based controller. In this study, $R_{\max}$ was set to a high value to study the demand for the thermal resistance upper limit of the AIS under different weather condition. During February 16th to February 17th, the interior AIS was maintained as low-level thermal resistance under MFORL control while the AIS under binary control switching to $R_{\max}$ might lead to overheating effect. Hence, RBE under MFORL control allows more flexible charging and discharging of the concrete layer. The thermal resistance variation range of AIS also indicates that the upper thermal resistance limit has an effect on the overall energy saving potential for different climate.

Table 2 presents the Long-term Percentage of Dissatisfied (LPD) in different scenarios during representative days shown in Figure 4 and Figure 5. RBEs under rule-based control and MFORL control provide 6% - 25% improvement for the LPD of baseline. The MFORL controller not only provide substantial AC load reduction but also improve thermal comfort as compared to the baseline in the case studies.

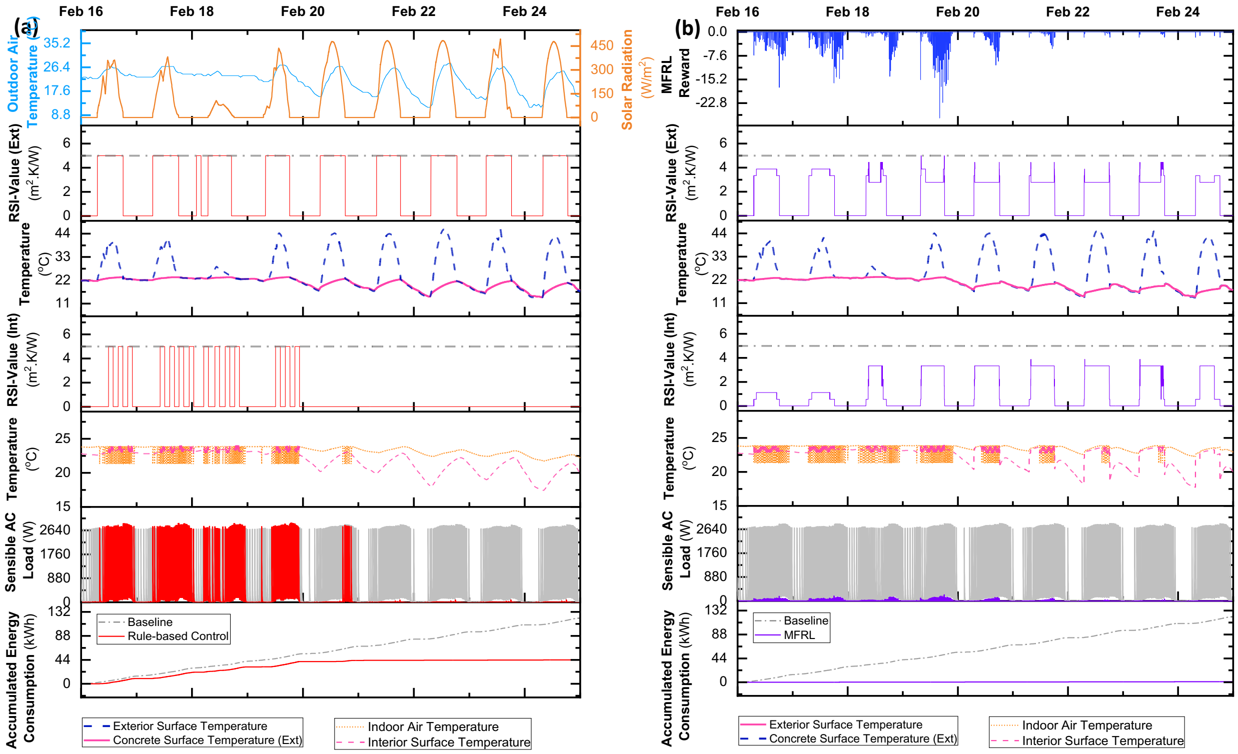

Figure 5 Simulation results for Miami FL with responsive building envelope (RBE): (a) rule-based control; (b) MFORL control

Table 2. LPD of the building in different scenarios

City	LPD (Baseline)	LPD (RBE under rule-based control)	LPD (RBE under MFORL control)
Miami, FL	7.91	7.72	7.43
Albuquerque, NM	8.16	5.90	6.10

CONCLUSION

Building energy simulations were performed on a representative residential thermal zone to test both a simple temperature-driven rule-based controller and a newly formulated model-free online reinforcement learning (MFORL) controller for responsive building envelopes (RBEs) that are consisted with a sensible thermal mass sandwiched between two active insulation layers (RSI = 0.01-5 m^2K/W). For the residential building modeled, it was found that under certain condition (i.e., when outdoor environment can provide beneficial heat flow to offset AC loads), the use of AIS can lead to significant reduction in AC loads. As for the two control methods tested for RBEs, the MFORL controller showed higher energy saving potential than the rule-based controller during the tested durations. Since MFORL does not rely on the formulation of a physics-based model, it has shown significant promise to be used for RBE control, especially under complex environmental conditions such as those during transitional seasons. Future work is needed to demonstrate the energy saving potential of MFORL in different thermal zones and under different climate conditions. Other performance metrics such as thermal comfort and grid flexibility can also be considered in the formulation of MFORL controllers.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support from U.S. National Science Foundation under grant CMMI-1954517.

NOMENCLATURE

C = Heat capacity of node

H = Heat transfer coefficient

LPD = Long-term Percentage of Dissatisfied

PPD = Predicted percentage of dissatisfied

Q = Heat rate

R = Thermal resistance

t = Time

T = Temperature

γ = weight for reward function

Subscripts

air_in = Indoor Air

i = Node sequence

max = Maximum

min = Minimum

Supercripts

j = Time step sequence

REFERENCES

Antretter, F. (2019) 'Active Insulation Systems', *BTO Peer Review Presentation*, pp. 1–21.

Antretter, F. and Boudreaux, P. (2019) 'Assessing the Potential of Active Insulation Systems to Reduce Energy Consumption and Enhance Electrical Grid Services', *2019 Buildings XIV International Conference*, pp. 12–15.

Baldi, S. *et al.* (2015) 'Model-based and model-free "plug-and-play" building energy efficient control', *Applied Energy*, 154, pp. 829–841. doi: 10.1016/j.apenergy.2015.05.081.

Benson, D. K., Potter, T. F. and Tracy, C. E. (1994) 'Design of a variable-conductance vacuum insulation', *SAE Technical Papers*, 103(1994), pp. 176–181. doi: 10.4271/940315.

Berge, A. *et al.* (2015) 'Effect from a Variable U-Value in Adaptive Building Components with Controlled Internal Air Pressure', *Energy Procedia*. Elsevier B.V., 78, pp. 376–381. doi: 10.1016/j.egypro.2015.11.677.

Biloria, N. and Sumini, V. (2007) 'Performative building skin systems: a morphogenomic approach towards developing real-time adaptive building skin systems', *International Journal of Architectural Computing*, 07(04), pp. 643–675. doi: 10.1260/1478-0771.7.4.643.

Brigham, J. (2015) *Collaborative Research : Adaptive and Reconfigurable Tiles for Building Surfaces*, *NSF*.

Carlucci, S. (2013) *Thermal Comfort Assessment of Buildings*. Springer Milan Heidelberg New York Dordrecht London. doi: 10.1007/978-88-470-5238-3.

Cui, H. and Overend, M. (2019) 'A review of heat transfer characteristics of switchable insulation technologies for thermally adaptive building envelopes', *Energy and Buildings*. Elsevier B.V., 199, pp. 427–444. doi: 10.1016/j.enbuild.2019.07.004.

Deardorff, J. W. (1978) 'Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation', *Journal of Geophysical Research*, 83(C4), p. 1889. doi: 10.1029/jc083ic04p01889.

He, Y. *et al.* (2020) 'Energy-saving potential of 3D printed concrete building with integrated living wall', *Energy and Buildings*. Elsevier B.V., 222, p. 110110. doi: 10.1016/j.enbuild.2020.110110.

ISO (2005) 'ISO 7730: Ergonomics of the thermal environment Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria', *Management*, 3, pp. 605–615. doi: 10.1016/j.soildyn.2004.11.005.

Jin, Q., Favoino, F. and Overend, M. (2017) 'Design and control optimisation of adaptive insulation systems for office buildings. Part 2: A parametric study for a temperate climate', *Energy*. Elsevier Ltd, 127, pp. 634–649. doi:

10.1016/j.energy.2017.03.096.
Joe, J. *et al.* (2013) 'Load characteristics and operation strategies of building integrated with multi-story double skin facade', *Energy and Buildings*. Elsevier B.V., 60, pp. 185–198. doi: 10.1016/j.enbuild.2013.01.015.
Kimber, M., Clark, W. W. and Schaefer, L. (2014) 'Conceptual analysis and design of a partitioned multifunctional smart insulation', *Applied Energy*. Elsevier Ltd, 114, pp. 310–319. doi: 10.1016/j.apenergy.2013.09.067.
Liu, S. and Henze, G. P. (2006) 'Experimental analysis of simulated reinforcement learning control for active and passive building thermal storage inventory: Part 2: Results and analysis', *Energy and Buildings*, 38(2), pp. 148–161. doi: 10.1016/j.enbuild.2005.06.001.
Loonen, R. C. G. M. *et al.* (2013) 'Climate adaptive building shells: State-of-the-art and future challenges', *Renewable and Sustainable Energy Reviews*, 25, pp. 483–493. doi: 10.1016/j.rser.2013.04.016.
Loonen, R. C. G. M. (2015) 'Bio-inspired Adaptive Building Skins', in Pacheco Torgal, F. et al. (eds) *Biotechnologies and Biomimetics for Civil Engineering*. Springer, pp. 115–134.
Loonen, R. C. G. M., Hoes, P. and Hensen, J. L. . (2014) 'Performance prediction of buildings with responsive building elements: Challenges and solutions', *Proceedings of building simulation and optimization*, (June), pp. 1–8. Available at: https://pure.tue.nl/ws/portalfiles/portal/3992582/392669136570421.pdf.
Michailidis, I. T. *et al.* (2018) 'Energy-efficient HVAC management using cooperative, self-trained, control agents: A real-life German building case study', *Applied Energy*. Elsevier, 211(October 2017), pp. 113–125. doi: 10.1016/j.apenergy.2017.11.046.
Mnih, V. *et al.* (2016) 'Asynchronous methods for deep reinforcement learning', *33rd International Conference on Machine Learning, ICML 2016*, 4, pp. 2850–2869.
Monetti, V. *et al.* (2015) 'Calibration of building energy simulation models based on optimization: A case study', *Energy Procedia*. Elsevier B.V., 78, pp. 2971–2976. doi: 10.1016/j.egypro.2015.11.693.
Mumme, S. and James, N. (2020) 'Smart and Efficient Building Envelopes : Thermal Switches and Thermal Storage for Energy Savings and Load Flexibility', *ASHRAE Transactions*, 126(2), pp. 140–148.
Park, B., Srubar, W. V. and Krarti, M. (2015) 'Energy performance analysis of variable thermal resistance envelopes in residential buildings', *Energy and Buildings*. Elsevier B.V., 103, pp. 317–325. doi: 10.1016/j.enbuild.2015.06.061.
Pflug, T. *et al.* (2015) 'Closed translucent façade elements with switchable U-value - A novel option for energy management via the facade', *Energy and Buildings*. Elsevier B.V., 86, pp. 66–73. doi: 10.1016/j.enbuild.2014.09.082.
Pflug, T. *et al.* (2018) 'Modeling of facade elements with switchable U-value', *Energy and Buildings*. Elsevier B.V., 164, pp. 1–13. doi: 10.1016/j.enbuild.2017.12.044.
Pinto, G. *et al.* (2021) 'Coordinated energy management for a cluster of buildings through deep reinforcement learning', *Energy*, 229. doi: 10.1016/j.energy.2021.120725.
Rupp, S. and Krarti, M. (2019) 'Analysis of multi-step control strategies for dynamic insulation systems', *Energy and Buildings*. Elsevier B.V., 204. doi: 10.1016/j.enbuild.2019.109459.
Rybkowski, Z. *et al.* (2015) *EAGER : Interaction of Smart Materials for Transparent , Self - regulating Building Skins*, *NSF*.
Schulman, J. *et al.* (2015) 'Trust region policy optimization', *32nd International Conference on Machine Learning, ICML 2015*, 3, pp. 1889–1897.
Schulman, J. *et al.* (2016) 'High-dimensional continuous control using generalized advantage estimation', *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, pp. 1–14.
Shekar, V. and Krarti, M. (2017) 'Control strategies for dynamic insulation materials applied to commercial buildings', *Energy and Buildings*. Elsevier B.V., 154, pp. 305–320. doi: 10.1016/j.enbuild.2017.08.084.
Stec, W. J. and Paassen, a. H. C. V (2005) 'Symbiosis of the double skin facade with the HVAC system', *Energy and Buildings*, 37(5), pp. 461–469. doi: 10.1016/j.enbuild.2004.08.007.
Sung, D. K. (2010) 'Skin Deep: Making Building Skins Breathe with Smart Thermo bimetals, Where Do You Stand', in *Proceedings of the 2011 ACSA National Conference*. Washington, DC: ACSA Press, pp. 145–152.
Tomko, J. A. *et al.* (2018) 'Tunable thermal transport and reversible thermal conductivity switching in topologically networked bio-inspired materials', *Nature Nanotechnology*. Springer US, 13(10), pp. 959–964. doi: 10.1038/s41565-018-0227-7.
Varga, S., Oliveira, A. C. and Afonso, C. F. (2002) 'Characterisation of thermal diode panels for use in the cooling season in buildings', *Energy and Buildings*, 34(3), pp. 227–235. doi: 10.1016/S0378-7788(01)00090-1.
Velikov K, T. G. (2013) 'Responsive Building Envelopes : Characteristics and Evolving Paradigms', in *Design and*

Construction of High-Performance Homes, pp. 75–92.
Wigginton, M. and Harris, J. (2002) *Intelligent Skins*, *Elsevier*.
Wu, Z. *et al.* (2014) 'a Comparative Study on Thermal Conductivity and Rheology Properties of Alumina and Multi-Walled Carbon Nanotube Nanofluids', *Frontiers in Heat and Mass Transfer*, 5(December). doi: 10.5098/hmt.5.18.
Yoon, Y. R. and Moon, H. J. (2019) 'Performance based thermal comfort control (PTCC) using deep reinforcement learning for space cooling', *Energy and Buildings*. Elsevier B.V., 203. doi: 10.1016/j.enbuild.2019.109420.
Zheng, Z. and Wang, N. (2002) 'Model-free control based on neural networks', *Proceedings of 2002 International Conference on Machine Learning and Cybernetics*, 4(November), pp. 2180–2183. doi: 10.1109/icmlc.2002.1175425.