This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

Optimal HVAC System Operation Using Online Learning of Interconnected Neural Networks

Ye-Eun Jang, *Student Member*, *IEEE*, Young-Jin Kim, *Senior Member*, *IEEE*, and João P. S. Catalão, *Senior Member*, *IEEE*

***Abstract*—Optimizing the operation of heating, ventilation, and air-conditioning (HVAC) systems is a challenging task that requires the modeling of complex nonlinear relationships among the HVAC load, indoor temperature, and outdoor environment. This paper proposes a new strategy for optimal operation of an HVAC system in a commercial building. The system for indoor temperature control is divided into three sub-systems, each of which is modeled using an artificial neural network (ANN). The ANNs are then interconnected and integrated into an optimization problem for temperature set-point scheduling. The problem is reformulated to determine the optimal set-points using a deterministic search algorithm. After the optimal scheduling has been initiated, the ANNs undergo online learning repeatedly, mitigating overfitting. Case studies are conducted to analyze the performance of the proposed strategy, compared to strategies with a pre-determined temperature set-point, an ideal physics-based building model, and other types of machine learning-based modeling and scheduling methods. The case study results confirm that the proposed strategy is effective in terms of the HVAC energy cost, practical applicability, and training data requirements.**

Index Terms—Artificial neural networks (ANNs); deterministic search; heating, ventilation, and air-conditioning (HVAC); online learning; temperature set-point scheduling.

NOMENCLATURE

The main notations used in this paper are summarized here.

***A. Sets and Indices*:**

d, t	indices for day and time
m, n	indices for neural networks and linear power blocks
max, min, ref, set	subscripts for maximum, minimum, reference, and set-point values
$e(\bullet, \bullet')$	normalized root mean square error between • and •′
$\mathbf{L}_1$, $\mathbf{L}_2$, $\mathbf{L}_3$	neural networks to model building thermal dynamics
$\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{S}_3$	original and reformulated optimization problems

***B. Parameters*:**

C^t	retail electricity price at time t
D_T	dead-band between T_{set}^t and T_i^t
$\mathbf{E}^t$	building thermal environments at time t
$F_{n,\tau}^t$	linear gradient of T_i^t at time t resulting from input power segment n of HVAC system at time τ
L_{P1}, L_{P2}, L_{P3}	maximum time delays of input data for neural networks
N_d	number of days for online supervised learning
N_{ET}, N_{EO}	numbers of epochs for network training and optimal scheduling
N_{HLm}	number of hidden layers of the m_{th} network
N_{HNm}	number of hidden nodes in each layer of the m_{th} network
N_{ID}	number of initial training datasets
N_S	number of linear power blocks
N_T	number of scheduling time intervals in a day
P_{max}, P_{min}	maximum and minimum power inputs of HVAC system
P_O	offset of reference power input to HVAC system
Q_i^t	internal thermal load of a building at time t
R_H, R_L	upward/downward ramp rate limits of the power input to HVAC system
R_T, R_O	learning rates for network training and optimal scheduling
T_a^t	adjacent room temperature at time t
T_e^t	evaporator-side air temperature at time t
T_{in}^t	indoor temperature at time t under the no cooling condition from 1 to t; i.e., the HVAC system remains off.
$T_{i,max}^t$, $T_{i,min}^t$	maximum and minimum limits of T_i^t at time t
$T_{set,max}$, $T_{set,min}$	maximum and minimum set-point temperatures
T_x^t	outdoor (condenser-side) air temperature at time t
k_P, k_I	proportional and integral gains of a thermostat controller
t_s, t_e	start- and end-times of the working hours in a building
Δt	unit time step
Δt_U	time period to update neural networks online
λ_y, λ_k, λ_h	weighting factors for HVAC energy cost and constraints for system controllable inputs and system states
$\delta_{n,max}$	maximum value of the n_{th} linear power block
ξ_T, ξ_O	weight decays for network training and optimal scheduling
μ_T, μ_O	weight change momentums for network training and optimal scheduling

***C. Variables*:**

C_E	daily energy cost of HVAC system
J	objective value of the reformulated optimization problem with penalty on the system inputs and states
P^t	power input to HVAC system at time t
P_c^t	power input to HVAC system at time t in a conventional strategy
P_{ref}^t	reference power input to HVAC system at time t
Q^t	cooling rate supplied by HVAC system at time t
T_{set}^t, T_i^t	set-point and indoor temperatures at time t
a_n^t	binary variables for piecewise linearization of the variation in T_i^t resulting from HVAC power segment n at time t
e^t	difference between T_{set}^t and T_i^t at time t
r_{CR}	ratio of reduction in daily energy cost of HVAC system
v_{TC}	sum of deviations of T_i^t from an acceptable range during a day
δ_n^t	input power assigned in the linear power block n at time t

Manuscript received September 6, 2020; revised December 20, 2020; accepted January 11, 2021. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1C1C1003361). (*Corresponding author: Y. Kim*).

Y. Jang and Y. Kim are with the Department of Electrical Engineering, Pohang University of Science and Technology (POSTECH), Pohang, Gyungbuk, 37673, Korea (e-mail: powersys@postech.ac.kr).

J. P. S. Catalão is with the Faculty of Engineering of the University of Porto, 4200-465 Porto, Portugal, and also with INESC TEC, 4200-465 Porto, Portugal (e-mail: catalao@fe.up.pt).

I. INTRODUCTION

COMMERCIAL buildings accounted for more than 36% of total energy consumption in the United States in 2019 [1]. Heating, ventilation, and air conditioning (HVAC) units represent approximately 40% of the electricity used in commercial buildings [2]. Therefore, significant attention has

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

been given to the modeling and optimal operation of HVAC systems to improve energy efficiency and reduce electricity bills in commercial buildings.

Physics-based modeling of HVAC units requires numerous parameters to reflect the complex nonlinear relationships among the HVAC load, indoor temperature, and outdoor environment [3]. Most of the physics-based modeling parameters are unknown and need to be extracted using sophisticated estimation techniques. Therefore, previous studies using simple RC circuit models need further analysis to reflect the thermal dynamics of buildings accurately [4]. Moreover, the types, sizes, and operating characteristics of HVAC units vary by manufacturer and by the building in which they are installed [5]. This prevents application of physics-based modeling and optimal operation to various buildings with different types of HVAC system.

To overcome these challenges, machine learning (ML) and artificial neural networks (ANNs) have increasingly been considered in recent studies on building energy management systems (BEMSs). This paper proposes a new ML-based strategy for an HVAC system in a commercial building, wherein the optimal temperature set-points are deterministically scheduled using the online supervised learning (SL) of interconnected ANNs. Specifically, the system for a building's temperature control is divided into three sub-systems: a thermostat controller, an HVAC unit, and a building envelope. Long short-term memory (LSTM) networks are implemented and trained to model each sub-system. The LSTM networks are then interconnected to establish a complete model of the temperature control system. Using the LSTM-based model, an optimization problem is formulated to schedule the optimal temperature set-points, given day-ahead forecasts of the electricity price and the thermal environment. The problem is then reformulated, so that the optimal solution can be deterministically searched for using a gradient descent (GD) algorithm. After the optimal scheduling has been initiated, the LSTM-based model continues to undergo online SL, as new data on the building's operation are collected. This gradually improves the accuracy of the LSTM-based model and hence the performance of the optimal scheduling. The results of sensitivity analyses and comparative case studies confirm that the proposed strategy ensures cost-effective operation of the HVAC system and the thermal comfort of occupants.

The main contributions of this paper are summarized below:

• To the best of our knowledge, this study is the first to develop and interconnect the ANN models of the sub-systems that are required for building temperature control, mitigating the complexity of the ANNs and improving the modeling accuracy of the building thermal dynamics.

• The interconnected ANNs are directly integrated into the optimization problem for temperature set-point scheduling. The problem is then reformulated to apply a deterministic search algorithm and find the optimal schedule within a reasonable computation time.

• The online SL is incorporated into the optimal scheduling of HVAC system operation, reducing the requirement for initial training data and hence facilitating the application of ML-based modeling and control in practice.

• The proposed strategy is comprehensively evaluated, both using sensitivity analyses and via comparison with strategies that use a traditional temperature setting rule, an ideal physics-based building model, and other types of ML algorithm.

II. Related Works

In recent years, Internet of things (IoT) technologies have been widely used to facilitate interactions between BEMSs and in-building infrastructures [6], [7], as significant attention has been given to improving building energy efficiency. The costs of IoT sensors and data analytics tools continue to decrease and they have become more widely and immediately available. Consequently, labeled datasets of HVAC unit operations and building thermal conditions have become increasingly available to BEMSs [7]–[9], enabling data-driven modeling and operation of building temperature control systems in practice.

Given this data availability, various ML algorithms have been used in recent studies on optimal control of indoor temperatures. For example, in [8]–[10], an ANN was trained offline via SL to model building thermal dynamics. Given the ANN model, the solution to the problem for the optimal HVAC system operation was searched for using heuristic algorithms, such as GA, PSO, and firefly algorithms. However, in a heuristic search, the solution is highly likely to fall into one of numerous local optima. Therefore, the optimization problem should be iteratively solved to find the best solution closer to the global optimum, increasing the computation time [8]. Moreover, in [8]–[10], only a single ANN was implemented to reflect the highly nonlinear characteristics of the building thermal dynamics. In practice, this risks compromising the modeling accuracy and hence the scheduling performance, even for the case of an ANN with deep hidden layers.

In [11]–[16], reinforcement learning (RL) was adopted to take advantage of the fact that it requires little knowledge of HVAC system operations and building thermal dynamics. For example, in [11] and [12], the optimal operation of air conditioners was explored using Deep Q-Network (DQN) and Deep Policy Gradient (DPG) algorithms. In [13]–[15], deep DPG (DDPG) and A3C algorithms were adopted to minimize HVAC energy consumption, considering the high-dimensional action spaces. In the RL algorithms, for each episode, an RL agent chooses an action based on the exploration-and-exploitation mechanism [16], where the agent explores untried actions to gain more experience and combines this with exploitation of the already known successful actions to obtain high long-term reward. In other words, the optimal HVAC load schedule still needs to be iteratively searched for using random variables. For the heuristic search, the number of learning episodes should be set to an arbitrarily high value, increasing the computation time.

In addition, the exploration-and-exploitation mechanism is highly likely to include the risk that a slight change in hyper-parameters for the RL agent's training can lead to unstable and poor control of the HVAC system, particularly in the initial learning episodes [17], [18]. When the spaces of states and actions are discretized, a large step size can also lead to poor capability of the RL agent to learn the problem characteristics

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

and failure to ensure the thermal comfort of occupants. This implies difficulties in directly applying RL-based control algorithms to real buildings where HVAC systems are currently in daily service. Therefore, in recent studies (e.g., [13] and [19]), data-driven models of HVAC systems and building envelopes were developed first, so that the RL agent was trained using the input and output datasets obtained from the models, as in the case of model predictive control (MPC) [20]. However, once the models are implemented, it can be more stable and time-efficient to apply SL-based control strategies using deterministic optimal solvers, rather than RL-based strategies.

The application of SL requires historical data on HVAC system operations under various building thermal environments. When the size of the historical dataset is small and the variability is limited, the ANNs are likely to be over-fitted [21]: i.e., too closely fitted to only a limited set of data points. The requirement for historical data needs to be mitigated for wide application of SL-based modeling and optimal operation. For example, in new buildings, insufficient historical data may have been collected. In traditional energy-inefficient buildings, a rule-based strategy is often adopted to operate HVAC systems with pre-determined temperature set-points. To reduce the data requirement, recent studies have been conducted on online SL. For example, in [22], the optimal operation of an air-conditioning system was achieved online, although variations in the ambient temperature and electricity price were not considered. In [23], hyper-parameters for optimal HVAC system operation were updated online; however, the temperature set-point was chosen from only a limited set of discrete values and was fixed during a day.

III. Modeling of Building Thermal Dynamics

A. ANN-based Modeling of Sub-systems

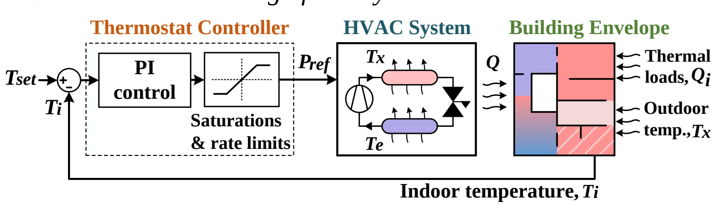

Fig. 1. A schematic diagram of a common system for building temperature control, consisting of a thermostat control loop, an HVAC unit, and a building envelope.

Fig. 1 shows a common system for indoor temperate control in a commercial building. It consists of three sub-systems: a thermostat-control loop, an HVAC unit, and a building envelope. Specifically, in the thermostat loop, a proportional-integral (PI) controller is adopted to adjust the reference power input P^t_{ref} of the HVAC unit, based on the difference between the set-point and actual values of the indoor temperature: i.e., T_{set}^t and T_i^t, respectively. In practice, the PI controller is accompanied by nonlinear signal processing functions, such as saturations and ramp rate limits, to ensure reliable system operation. The HVAC unit receives P^t_{ref} as an input signal and provides thermal energy Q^t to the envelope, given the ambient temperature T_x^t and the evaporator-side air or water temperature T_e^t. In this paper, a variable speed heat pump is considered as an example of an HVAC unit [8]. The time response of the variable speed drive is fast and, consequently, the actual power input P^t is almost the same as P_{ref}^t (i.e., $P^t \approx P_{ref}^t$), particularly in the scheduling time horizon. In the building envelope, the profile of T_i^t is determined by the HVAC system operation (i.e., Q^t) and the building thermal environments $\mathbf{E}^t$, such as T_x^t, T_e^t, and indoor thermal load Q_i^t.

Each sub-system is modeled using an ANN, as shown in Fig. 2. The ANNs are then linked together, based on the inter-connections of the sub-systems, as discussed above. The operating characteristics of each sub-system can successfully be reflected into an ANN with a rather simple architecture. This mitigates the overall complexity of the ANN model that represents the complete system for the building temperature control, shown in Fig. 1. By contrast, the conventional modeling methods often consider only a single ANN [8]–[10]. The ANN then needs to be significantly complicated and deep to reflect the operation of the complete system accurately, requiring a large amount of building operation data. This implies the risk of compromising modeling accuracy and hence the temperature control performance for a practical case with data of limited size and variability.

B. ANN Architecture and Training

For the sub-systems, the ANNs are implemented in the form of an LSTM network, which is widely used for time-series data learning and system identification. Note that the proposed strategy can readily be achieved using different types of ANNs, as discussed in Section V-D. Specifically, the LSTMs consist of multiple hidden layers, each of which includes multiple hidden nodes with self-loops. Furthermore, Fig. 2 shows that the LSTM $\mathbf{L}_1$ has an inner feedback loop between the output and input neurons for P^t, which is indicated by the red circles.

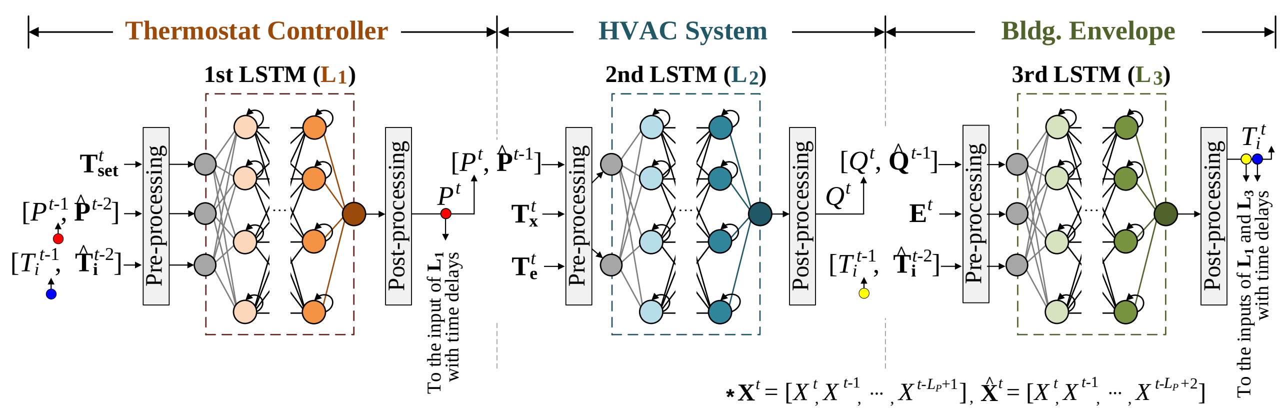

Fig. 2. Interconnection of the LSTM networks that correspond to the models of the thermostat controller, HVAC system, and building envelope, respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

Similarly, $\mathbf{L_3}$ has an inner feedback loop of T_i^t, marked by the yellow circles. An outer feedback loop of T_i^t also exists between the output neuron of $\mathbf{L_3}$ and the input neuron of $\mathbf{L_1}$, which is represented by the blue circles. Moreover, $\mathbf{L_{1\text{–}3}}$ have pre- and post-processors to normalize the input data and recover the output data with their original units, respectively, preventing the training speed from dropping too low.

In addition, each LSTM has a single output neuron and multiple input neurons. As shown in Fig. 2, the outputs of $\mathbf{L_{1\text{–}3}}$ are defined as P^t, Q^t, and T_i^t, respectively. The inputs of $\mathbf{L_1}$ are the current and time-delayed values of T_{set}^t and the time-delayed values of P^t and T_i^t. For $\mathbf{L_2}$, the inputs are the current and time-delayed values of P^t, T_x^t, and T_e^t. The inputs of $\mathbf{L_3}$ are set to the current and time-delayed values of Q^t and $\mathbf{E}^t$ and the time-delayed T_i^t. In this study, the time-delayed inputs of $\mathbf{L_{1\text{–}3}}$ are explicitly considered to achieve better accuracy in modeling the building thermal dynamics by reflecting the effects of the integral controller in the thermostat loop, the heat exchanger in the HVAC system, and the thermal energy storage inherent in the building envelope, respectively. Specifically, the search range for the hyper-parameters of $\mathbf{L_{1\text{–}3}}$ is established by the minimum and maximum values of the time delays of input neurons, the numbers of hidden layers and neurons, and the learning rates, considering the trade-off between the modeling accuracy and the computational burden. While examining all possible combinations, one is selected that leads to good training and testing results for historical datasets. Through this procedure, the maximum time delays of the inputs of $\mathbf{L_{1\text{–}3}}$ are set to L_{P1} = 24 h, L_{P2} = 4 h, and L_{P3} = 4 h, respectively. For brevity, each LSTM has the same values of L_P for its inputs, and the selected hyper-parameters of $\mathbf{L_{1\text{–}3}}$ are fixed during the online SL, as discussed in Section IV-B.

The individual LSTMs are trained separately using the database of a BEMS to determine the weighting coefficients and biases for all the input, hidden, and output neurons. The separate training can reduce the structural complexity of the LSTMs, facilitating modeling of the temperature control system. The feedback loops for each LSTM are also open, so that the actual time-delayed data can be fed into the input neurons and hence an SL algorithm can be applied for the LSTM training. The training data are obtained during the actual, normal operation of the temperature control system, ensuring the modeling convergence of $\mathbf{L_{1\text{–}3}}$. Moreover, the physics-based modeling parameters of the HVAC system and building envelope are not required to train the LSTMs and hence formulate the optimization problem, discussed in Section V, wherein the LSTMs are integrated for optimal scheduling of the set-point temperatures. This enables wide application of the proposed strategy in practical BEMSs. After the training, the LSTMs are then interconnected and tested with closed feedback loops, so that the outputs estimated at the current time step can be used as the time-delayed inputs at the next step. This also enables the interconnected LSTMs to reflect the interactions among the sub-systems and hence the operating characteristics of the completed system.

IV. Optimal Scheduling Integrated with Online SL

In the proposed strategy, the optimal operation of the HVAC system is scheduled for the next 24 hours, based on day-ahead forecasts of the electricity price and building thermal conditions. The scheduling is consistent with current practices for demand response (DR) [24] and existing strategies for scheduling of power system operation [25], [26]. Numerous forecasting algorithms have been discussed, for example, in [27]–[29] and, therefore, appropriate algorithms can readily be selected and incorporated into the proposed strategy. In this study, the forecast data are assumed to be already available in the BEMS database for brevity, as in [8] and [9]; integration has been left for future research.

A. Optimization Problem Formulation

Using the trained $\mathbf{L_{1\text{–}3}}$, the optimal schedule for T_{set}^t can be determined by solving $\mathbf{S_1}$ as:

$\mathbf{S_1}$: Problem for optimal HVAC system operation

$$\underset{T_{set}^t}{\arg\min} \quad C_E = \sum_{t=1}^{N_T} C^t P^t, \tag{1}$$

subject to

$$T_{set,\min} \le T_{set}^t \le T_{set,\max}, \quad \forall t, \tag{2}$$

$$T_{i,\min}^t \le T_i^t \le T_{i,\max}^t, \quad \forall t, \tag{3}$$

$$P_{\min} \le P^t \le P_{\max}, \quad \forall t, \tag{4}$$

$$R_L \le \left(P^t - P^{t-\Delta t}\right) / \Delta t \le R_H, \quad \forall t, \tag{5}$$

where

$$P^t = \mathbf{L_1}\left(T_{set}^t, \cdots, T_{set}^{t-L_{P1}+1}, P^{t-1}, \cdots, P^{t-L_{P1}}, T_i^{t-1}, \cdots, T_i^{t-L_{P1}}\right), \quad \forall t, \tag{6}$$

$$Q^t = \mathbf{L_2}\left(P^t, \cdots, P^{t-L_{P2}+1}, T_x^t, \cdots, T_x^{t-L_{P2}+1}, T_e^t, \cdots, T_e^{t-L_{P2}+1}\right), \quad \forall t, \tag{7}$$

$$T_i^t = \mathbf{L_3}\left(Q^t, \cdots, Q^{t-L_{P3}+1}, \mathbf{E}^t, \cdots, \mathbf{E}^{t-L_{P3}+1}, T_i^{t-1}, \cdots, T_i^{t-L_{P3}}\right), \quad \forall t. \tag{8}$$

The objective function (1) aims to minimize the energy cost C_E of the HVAC system: i.e., the 24-h sum of the hourly-varying retail electricity price C^t multiplied by the power input P^t of the HVAC system. Note that C^t can be negative, for example, when there is an excess of renewable generation [30].

In the set of constraints, (2) shows the limits of the operating range of the thermostat controller (i.e., from $T_{set,min}$ = 15ºC to $T_{set,max}$ = 35ºC) to secure reliable operation of the HVAC system. Moreover, (3) represents that T_i^t should be maintained within an acceptable range from $T_{i,min}^t$ to $T_{i,max}^t$ to ensure the thermal comfort of occupants. Note that T_{set}^t and T_i^t can differ under normal operating conditions of the HVAC system, mainly due to the large thermal capacity of the building envelope. The constraints (4) require P^t to be maintained between P_{max} and P_{min}; in this paper, these are set to the rated power input and zero, respectively. Furthermore, (5) specifies the limits on the upward and downward ramp rates of P^t for the time period Δt = 1 h. In (5), P^t at t = 0 h is set to zero, assuming that the HVAC system is turned off at night (after 7 pm to midnight) when the commercial building has low occupancy.

In (6)–(8), the LSTM-based sub-system models, discussed in Section III, are parameterized as the functions $\mathbf{L_{1\text{–}3}}(\cdot)$, in which the current and time-delayed inputs and the output are specified. In other words, the operating characteristics of the sub-systems are integrated as nonlinear equality constraints in $\mathbf{S_1}$, so that the optimal solution of $\mathbf{S_1}$ reflects the relationships between the controllable variable T_{set}^t and the dependent variables P^t, Q^t, and T_i^t, given the constant vector $\mathbf{E}^t$, for the current and delayed time steps. Specifically, in (6), $\mathbf{L_1}(\cdot)$ specifies the relationships between the input variables T_{set}^t and T_i^t and the output variable P^t of the thermostat controller, thus establishing the links of (2) and (3) with (4) and (5). Similarly, in (7) and (8), $\mathbf{L_2}(\cdot)$ and $\mathbf{L_3}(\cdot)$ connect the variable P^t in (4) and (5) with the variables Q^t and hence T_i^t in (3).

The optimization problem $\mathbf{S_1}$ [i.e., (1)–(8)] can be equivalently expressed in a compact form using the simple expressions of T_{set}^t, P^t, Q^t, T_i^t, and $\mathbf{E}^t$, as well as of $\mathbf{L_{1\text{–}3}}(\cdot)$, as:

$\mathbf{S_2}$: Compact form of the original problem $\mathbf{S_1}$

$$\arg\min_{u^t} \quad C_E = \sum_{t=1}^{N_T} y^t, \tag{9}$$

$$\text{subject to} \quad u_{\min} \le u^t \le u_{\max}, \quad \forall t, \tag{10}$$

$$\mathbf{s}_{\mathbf{min}}^t \le \mathbf{s}^t \le \mathbf{s}_{\mathbf{max}}^t, \quad \forall t, \tag{11}$$

$$\mathbf{s}^t = \boldsymbol{g}(\mathbf{s}^{t-1}, \mathbf{v}^t), \quad \forall t, \tag{12}$$

$$y^t = f(\mathbf{s}^t, \mathbf{v}^t), \quad \forall t. \tag{13}$$

In $\mathbf{S_2}$, y^t and u^t are defined as the output $C^t{\cdot}P^t$ and the controllable input T_{set}^t, respectively, of the system for optimal building temperature control. Moreover, for notational simplicity, a vector $\mathbf{v}^t$ is used to represent the system inputs $[u^t, \mathbf{w}^t]^{\mathrm{T}}$, including the system disturbances $\mathbf{w}^t = \mathbf{E}^t$. Similarly, $\mathbf{s}^t$ is used to indicate the system states $[T_i^t, Q^t, P^t, \Delta P^t]^{\mathrm{T}}$ that characterize the operating condition of the temperature control system at time t. Then, (1) can be equivalently represented as (9). Moreover, (2)–(5) can be simply expressed as (10) and (11), and (6)–(8) correspond to (12), where $\boldsymbol{g}(\cdot)$ represents a set of nonlinear functions. Therefore, $\mathbf{S_1}$ and $\mathbf{S_2}$ are the same as each other. To complete the $\mathbf{S_2}$ formulation using the simplified notation, (13) is added to connect y^t with $\mathbf{s}^t$ and $\mathbf{v}^t$, considering the relationship of P^t with T_i^t, T_{set}^t, and $\mathbf{E}^t$.

As clearly shown in (9)–(13), optimal operation of the HVAC system is achieved by solving a constrained nonlinear optimization problem. To find the optimal u^t, $\mathbf{S_2}$ is relaxed to an unconstrained problem using the continuous, quadratic penalty functions of (10) and (11) as:

$\mathbf{S_3}$: Reformulated problem of $\mathbf{S_2}$

$$\arg\min_{u^t} \sum_{t=1}^{N_T} \left\{ \lambda_y \left(y^t\right)^2 \cdot \operatorname{sgn}\left(y^t\right) + \lambda_k \left(k^t\right)^2 + \lambda_h \left(\mathbf{h}^t\right)^T \cdot \mathbf{h}^t \right\}, \tag{14}$$

where

$$k^t = \begin{cases} u^t - u_{\max} & \text{for } u^t > u_{\max} \\ u_{\min} - u^t & \text{for } u^t < u_{\min} \\ 0 & \text{otherwise} \end{cases}, \quad \forall t, \tag{15}$$

$$\mathbf{h}^t = \begin{cases} \mathbf{s}^t - \mathbf{s}_{\mathbf{max}}^t & \text{for } \mathbf{s}^t > \mathbf{s}_{\mathbf{max}}^t \\ \mathbf{s}_{\mathbf{min}}^t - \mathbf{s}^t & \text{for } \mathbf{s}^t < \mathbf{s}_{\mathbf{min}}^t \\ 0 & \text{otherwise} \end{cases}, \quad \forall t, \tag{16}$$

(12) and (13).

Specifically, in (14), the objective function is implemented using a quadratic function of the HVAC energy cost (i.e., $(y^t)^2 = (C^t{\cdot}P^t)^2$); $\operatorname{sgn}(y^t)$ is also taken into account, because C^t can be negative. Moreover, (10) and (11) are relaxed to (15) and (16), respectively, and then added to (14) in a quadratic form of the penalties incurred when the constraints on u^t and $\mathbf{s}^t$ are violated. In other words, $\mathbf{S_3}$ still reflects the operational constraints of the sub-systems. In (14), λ_k and λ_h are the corresponding penalty factors. Large values of the penalty factors lead to good consistency between the optimal solutions of $\mathbf{S_2}$ and $\mathbf{S_3}$; see Appendix A. Note that penalty factors that are too large are likely to create steep valleys on the constraint boundaries, rendering it difficult to solve $\mathbf{S_3}$ within a reasonable computation time. Therefore, it is common to apply penalty factors with small values and gradually increase them [31].

A GD algorithm [32], [33] is adopted to search for the minimum of the continuous, nonlinear function $\mathbf{S_3}$, where the next step is determined proportional to the negative gradient of $\mathbf{S_3}$ at the current step. Since this requires only the first derivative, the GD solver can readily be implemented in the BEMS, facilitating the optimal operation of the HVAC system in practice. Moreover, unlike heuristic, RL-based algorithms, the GD solver is deterministic and hence ensures that the optimal solution of $\mathbf{S_3}$ leads to stable, reliable system operation.

B. Online Supervised Learning

After optimal scheduling of T_{set}^t has been initiated, $\mathbf{L_{1\text{–}3}}$ undergo repeated online SL, as new data of T_{set}^t, P^t, Q^t, and T_i^t are obtained for various profiles of C^t and $\mathbf{E}^t$. This gradually mitigates the overfitting of $\mathbf{L_{1\text{–}3}}$. In other words, $\mathbf{L_{1\text{–}3}}$ become well adapted to changes in the operating conditions of the building, further improving the accuracy of modeling the building thermal dynamics. Specifically, Fig. 3 shows a flowchart for the online SL of $\mathbf{L_{1\text{–}3}}$. In Step 1, the optimal day-ahead scheduling of T_{set}^t is initiated, after $\mathbf{L_{1\text{–}3}}$ are trained with the initial historical data of the BEMS. Due to the small size and variability of the data, $\mathbf{L_{1\text{–}3}}$ are likely to be rather inaccurate, limiting the performance of the optimal scheduling. In Step 2, the HVAC system operates according to the optimal schedule of T_{set}^t on day d, and the BEMS collects the corresponding dataset $[T_{set}^t, P^t, Q^t, \mathbf{E}^t, T_i^t]$ for $1 \le t \le N_T$. The profiles of the dataset agre likely to differ from those of the historical BEMS datasets before the optimal scheduling is initiated. This increases the variability in the training data, improving the accuracy of $\mathbf{L_{1\text{–}3}}$ when they are re-trained using the newly collected dataset in Step 3. The re-training is

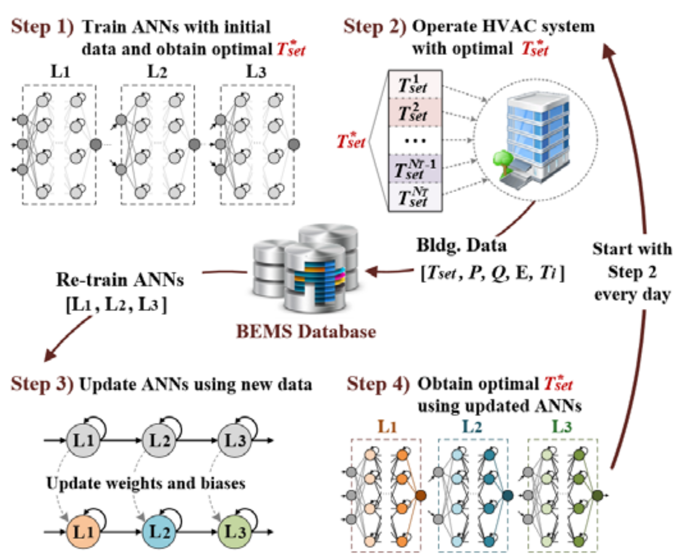

Fig. 3. Flowchart for the online SL of $\mathbf{L_{1\text{–}3}}$ that is integrated with the optimal scheduling of the HVAC system.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

conducted using the historical and online datasets for a number of epochs, and stops when the modeling accuracy at the current epoch is marginally improved, compared to that at previous epochs. In Step 4, the optimization problem $\mathbf{S_3}$ is updated using the retrained $\mathbf{L_{1–3}}$ and solved for the forecasts of $\mathbf{E}^t$ on day d+1. The improved accuracy of $\mathbf{L_{1–3}}$ will lead to expansion of the feasible solution area of $\mathbf{S_3}$, enhancing the performance of the optimal HVAC system operation. Step 2 is then repeated on day d+1 with the new optimal schedule of T_{set}^t. In this paper, Steps 2–4 are performed on each scheduling day during the period from day d = 1 to N_d to achieve continual improvement of the modeling accuracy and the scheduling performance. In practice, Steps 2–4 can be conducted once every several days and repeated continuously until the results are satisfactory.

V. Case Studies and Simulation Results

A. Test Conditions

The proposed strategy was tested for an experimental setup of an office building with an HVAC system, as shown in Fig. 4. Briefly, the experimental setup is divided into test and climate rooms, both of which are within a larger laboratory room with a temperature of T_a^t. The test room has lights and heat sources to emulate the internal thermal load Q_i^t of a common office. The walls and floor consist of multiple layers of different building materials. For the case studies, the power rating of the HVAC system was set to P_{max} = 50 kW, and a scaled fraction of the corresponding Q^t was used to control T_i^t in the test room. The climate room contains a separate heating unit to emulate the building thermal environments $\mathbf{E}^t$. For the experimental setup, a building simulator was implemented in [34] to estimate T_i^t for P^t, given $\mathbf{E^t} = [T_x^t, T_a^t, T_e^t, Q_i^t]$. In this study, the simulator was further extended by integrating the thermostat control loop with the HVAC system, as shown in Fig. 1; this enabled indirect control of the HVAC unit, as is common in real buildings.

To establish the initial training datasets, the building simulator was run using the data of Q_i^t estimated from a real building [34], [35] and of T_x^t measured in Boston from June 1 to August 31 of 2017–2019 [36], as shown in Fig. 5(a) and (b), respectively. Note that Q_i^t also can be surveyed and measured for benchmark buildings [37]. Given Q_i^t and T_x^t, the simulator was run with the pre-determined profiles of T_{set}^t, such that T_i^t was controlled within an acceptable range under the conditions of traditional HVAC system operation. Fig. 5(c) shows the corresponding profiles of P^t obtained from the simulation runs. Moreover, Fig. 5(d) shows the profiles of C^t [38] for the same time period as when the Q_i^t and T_x^t data were acquired. Note that on several days, C^t decreased below zero in the early morning. The sizes of the initial datasets [T_{set}^t, P^t, Q^t, $\mathbf{E}^t$, T_i^t] were 1,200 (i.e., 50 days) and 8 with respect to time and objects, respectively. The size with respect to time continued to increase, as the optimal profiles of T_{set}^t, P^t, Q^t, and T_i^t were obtained from $\mathbf{S_3}$ for days d = 1 to N_d (i.e., 200), as discussed in Section IV-B. In other words, the online SL was conducted while training $\mathbf{L_{1–3}}$ and solving $\mathbf{S_3}$ during the period from d = 1 to N_d. Note that the time-delayed data for the objects were not considered in the size estimation. The datasets were then randomly shuffled and divided into three parts with the ratios of 0.8:0.1:0.1 for the training, validation, and testing, respectively.

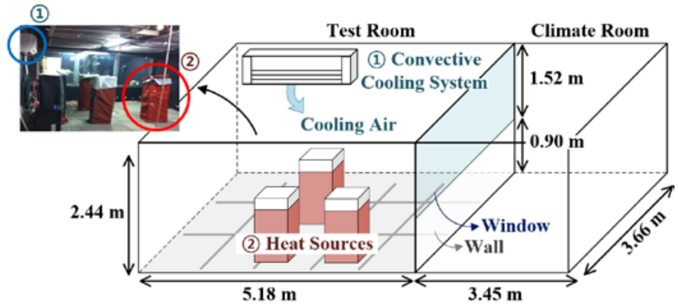

Fig. 4. Experimental setup for the data acquisition and comparative case studies

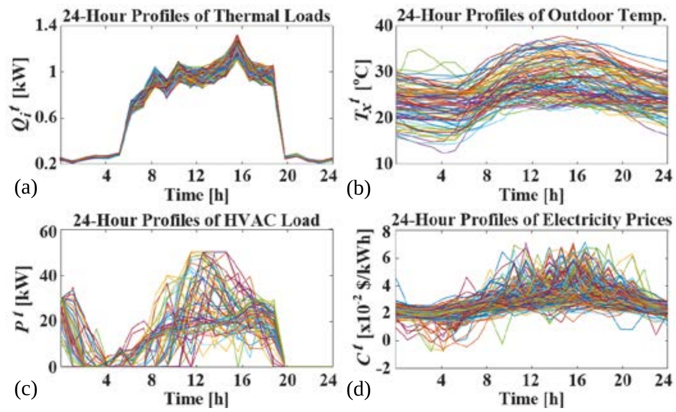

Fig. 5. Case study conditions from June 1 to August 31, 2017: (a) Q_i^t, (b) T_x^t, (c) P^t, and (d) C_t. The profiles of Q_i^t, T_x^t, P^t, and C^t in 2018 and 2019 were similar.

Table I. Parameter Values for the Case Studies

Modeling and scheduling	Parameters	Values	Units
Scheduling conditions	N_T	24	[h]
	$\Delta t, \Delta t_U$	1, 24	
	N_d	200	[d]
	N_{ID}	1200	[h]
	$\lambda_y, \lambda_k, \lambda_h$	2.0, 0.5, 7.5	
Thermostat controller	$T_{set,min}$, $T_{set,max}$	15, 35	[ºC]
	D_T	1	[h]
HVAC system	P_{min}, P_{max}	0, 50	[kW]
	R_L, R_H	–40, 30	
Building envelope	$T_{i,min}^t$, $T_{i,max}^t$	22, 24	[ºC]
	t_s, t_e	7, 19	[h]
Physics-based modeling	N_S	4	
LSTM-based modeling	L_{P1}, L_{P2}, L_{P3}	24, 4, 4	[h]
	N_{HL1}, N_{HL2}, N_{HL3}	3, 4, 3	
	N_{HN1}, N_{HN2}, N_{HN3}	15, 20, 20	
	N_{ET}, N_{EO}	5000, 1000	
	R_T	4×10^{-3}	
	R_{O1}, R_{O2}	10^{-3}, 10^{-4}	
	ξ_T, ξ_O	0, 0.3	
	μ_T, μ_O	0, 0.3	

Table II. Features of the Proposed, Ideal, and Rule-based Strategies

Strategies		Set-point temperatures	Building modeling	Opt. solver	Online learning
Proposed	Case 1	optimized	actual data of building operation	GD	○
Ideal	Case 2	optimized	fully-informed model parameters	MILP	-
Rule-based	Case 3	pre-determined	-	-	-

Table I lists the parameter values used for the modeling and optimal operation of the HVAC system in the case studies. The parameter values were determined mainly based on [8]–[10] and considering the current practices for DR, the sampling rates

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

of the BEMS datasets, and the convergence rates of the solution of $\mathbf{S}_3$. In particular, the learning rate for the LSTM training was set to be a small value of 4×10^{-3}, and the learning rate for the GD solver was reduced from $R_{O1} = 10^{-3}$ to $R_{O2} = 10^{-4}$ when the epoch number increased to greater than two thirds of the total number of epochs. This aimed to achieve high accuracy of the LSTM models $\mathbf{L}_{1\text{–}3}$ and the optimal solution to $\mathbf{S}_3$. Moreover, λ_y and λ_h were set to relatively large values to reduce the HVAC energy cost while ensuring the occupants' thermal comfort. By contrast, λ_k did not have to be set to a large value, because T_{set}^t and P^t varied within the acceptable ranges before the proposed strategy was applied. In other words, $\mathbf{L}_1$ and $\mathbf{L}_2$ were trained with the historical datasets of T_{set}^t and P^t ranging only between $T_{set,min}$ and $T_{set,max}$ and between P_{min} and P_{max}, respectively. For simplicity, λ_y, λ_k, and λ_h were fixed during $1 \le d \le N_d$.

The HVAC system operations were compared for three cases: the proposed SL-based strategy (Case 1), an ideal physics-based strategy (Case 2), and a traditional rule-based strategy (Case 3). Table II lists the main features of Cases 1–3. In Case 2, the piecewise linear equations for variations in T_i^t for a change in P^t were established using the complete information on the physics-based modeling parameters of the HVAC system and building envelope [39]; see (B1)–(B5) in Appendix B. The optimal schedule of T_{set}^t was then obtained by replacing (6)–(8) with (B1)–(B5) and then applying mixed- integer linear programming (MILP). Note that Case 2 is referred to as the ideal case, because most of the information is not available in practice. In Case 3, T_{set}^t was fixed at 23°C, regardless of the variation in C^t. For fair comparison of Cases 1–3, the HVAC system was assumed to be capable of operating from t = 1 h in the case studies. This also allowed the building to take advantage of pre-cooling for all Cases 1–3; late start of HVAC operation has the risk of causing an increase in C_E and a deviation of T_i^t from the acceptable range.

B. Improvement via Online Supervised Learning

The accuracy of the LSTM-based building model was verified by comparing the actual values of P^t, Q^t, and T_i^t in the testing datasets (discussed in Section V-A) with the corresponding estimates obtained from $\mathbf{L}_{1\text{–}3}$. Note that the estimates were acquired after training and interconnecting $\mathbf{L}_{1\text{–}3}$. Fig. 6 shows the results of the comparisons for d = 1 and N_d, where the x- and y-axes represent the actual values and the estimates, respectively. For d = 1, the normalized root mean square errors (nRMSEs) of $\mathbf{L}_{1\text{–}3}$ were estimated to be rather considerable: i.e., 1.1×10^{-1}, 9.5×10^{-3}, and 5.8×10^{-3}, respectively. As the online SL and optimal scheduling continued, the nRMSEs for $d = N_d$ were reduced to low levels of 9.1×10^{-3}, 3.1×10^{-3}, and 4.8×10^{-4}, respectively. Fig. 7 shows the variations in the nRMSEs over the period from d = 1 to N_d. For all $\mathbf{L}_{1\text{–}3}$, the nRMSEs were reduced rapidly during the initial period and decreased gradually for the remaining period. The results of the case studies confirmed that the online SL integrated with the optimal scheduling is effective in improving the accuracy of LSTM-based models of sub-systems (and hence the complete system) for building temperature control. In particular, the reduction of the nRMSEs for the testing datasets verified not only the improvement of the modeling accuracy of $\mathbf{L}_{1\text{–}3}$ but also

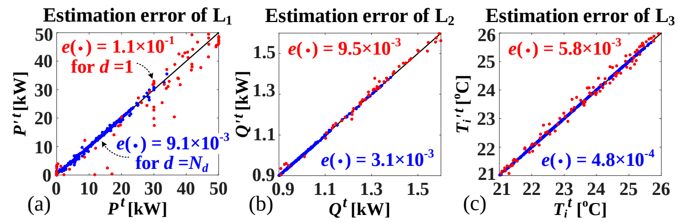

Fig. 6. Comparisons of the actual and estimated values of (a) P^t, (b) Q^t, and (c) T_i^t. The red and blue dots indicate the test results for d = 1 and N_d, respectively.

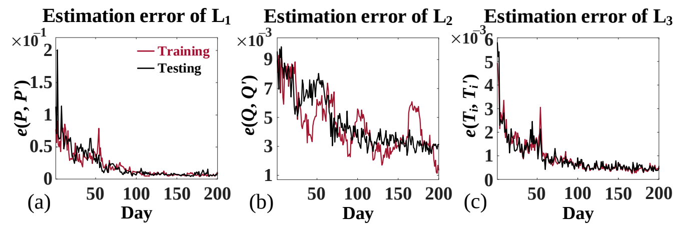

Fig. 7. Variations in the nRMSEs of (a)–(c) $\mathbf{L}_{1\text{–}3}$ during the online SL.

TABLE III. AVERAGE VALUES OF THE COST REDUCTION RATES FOR CASE 1

Periods	$0 < d \le N_d/4$	$N_d/4 < d \le N_d/2$	$N_d/2 < d \le 3N_d/4$	$3N_d/4 < d \le N_d$
r_{CR}^{avg} [%]	20.96	21.66	22.62	23.36

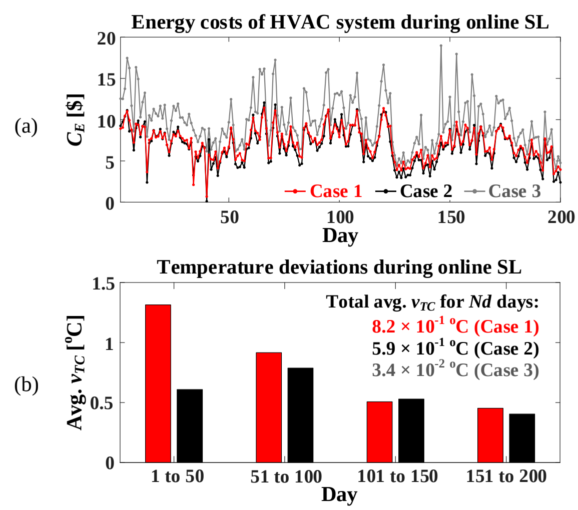

Fig. 8. Optimal scheduling results for Cases 1–3 over the time period from d = 1 to N_d: (a) C_E and (b) the average of v_{TC}.

the enhancement of their generalization capability, because the testing datasets were not used to train $\mathbf{L}_{1\text{–}3}$, as discussed in Section V-A. In other words, as the online SL continued, $\mathbf{L}_{1\text{–}3}$ became less over-fitted and hence more capable of accurately predicting the outputs of the sub-systems for the unseen inputs. Figs. 6 and 7 show that the nRMSEs of $\mathbf{L}_1$ were estimated to be higher than those of $\mathbf{L}_2$ and $\mathbf{L}_3$. This was mainly because the output of $\mathbf{L}_1$ (i.e., $P^t \approx P_{ref}^t$) changed faster and with larger magnitudes than the outputs of $\mathbf{L}_2$ and $\mathbf{L}_3$ (i.e., Q^t and T_i^t, respectively) due to the thermal capacity inherent in the HVAC refrigerant loop and building envelope. Note that in the case studies, the cooling energy supplied by the HVAC system was assumed to be equally divided into Q^t and used to control T_i^t in the test building room, as discussed in Section V-A.

In addition, Table III and Fig. 8 show the optimal scheduling results for the proposed strategy (i.e., Case 1) in comparison with those for the ideal and traditional strategies (i.e., Cases 2 and 3). For Cases 1 and 3, C_E was calculated as $\Sigma_t C^t \cdot P^t$ and $\Sigma_t C^t \cdot P_c^t$, respectively, during $1 \le t \le N_T$. The cost reduction rate

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

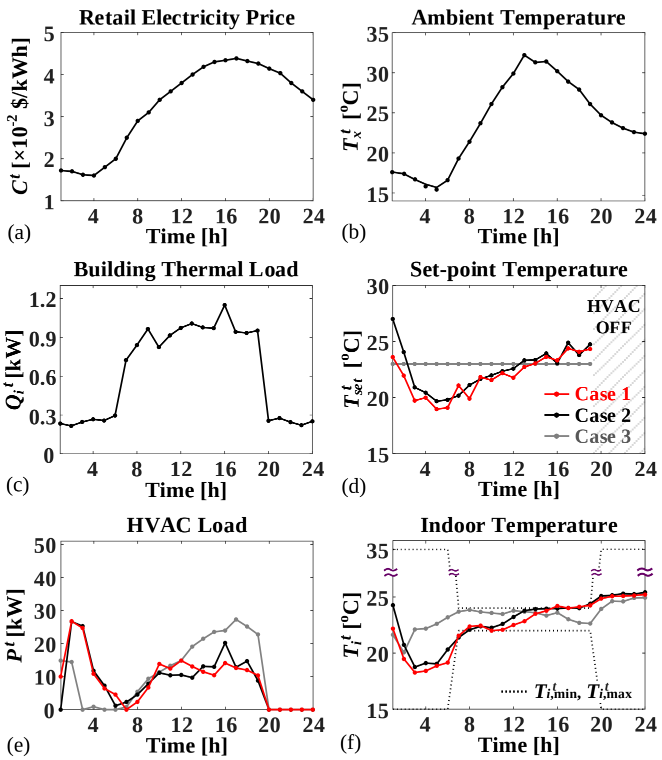

Fig. 9. Comparisons of the 24-h schedules for Cases 1–3: (a) C^t, (b) T_x^t, (c) Q_i^t, (d) T_{set}^t, (e) P^t, and (f) T_i^t.

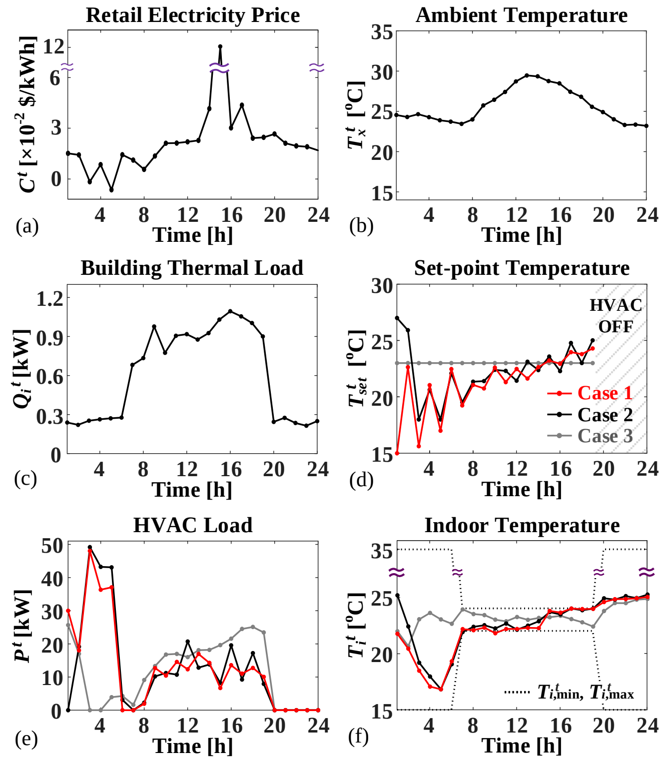

Fig. 10. Comparisons of the 24-h schedules for Cases 1–3 for different profiles of C^t and $\mathbf{E}^t$: (a) C^t, (b) T_x^t, (c) Q_i^t, (d) T_{set}^t, (e) P^t, and (f) T_i^t. In Fig. 10(a), the y-axis is broken to accommodate the peak of C^t for t = 15 h.

TABLE IV. COMPARISONS OF THE PROPOSED STRATEGY WITH THE IDEAL AND RULE-BASED STRATEGIES

Profiles of C^t and $\mathbf{E}^t$		Proposed (Case 1)	Ideal (Case 2)	Rule-based (Case 3)
Fig. 9	C_E [\$]	6.75	6.74	9.39
	r_{CR} [%]	28.1	28.2	-
	v_{TC} [°C]	1.03	1.03	0
Fig. 10	C_E [\$]	4.90	4.84	8.41
	r_{CR} [%]	41.7	42.4	-
	v_{TC} [°C]	0.18	0.08	0

r_{CR}^d for day d is then estimated as:

$$r_{CR}^d := \sum_{t=1}^{N_T} C^t \left(P^t - P_c^t\right) \Big/ \sum_{t=1}^{N_T} C^t P_c^t \quad . \tag{17}$$

In Table III, the average value of r_{CR}^d during the period of every $N_d/4$ (i.e., 50) days is calculated as:

$$r_{CR}^{avg} := \frac{1}{N_d/4} \sum_{d \in \mathbf{\Gamma}_i} r_{CR}^d \ , \tag{18}$$

where $\mathbf{\Gamma}_i = \{(i-1)\cdot N_d/4+1, \cdots, i\cdot N_d/4\}$ for $i = 1, \cdots, 4$. As the online SL continued, the average reduction rate r_{CR}^{avg} gradually increased from 20.96% to 23.36%. Fig. 8(a) shows the comparisons of C_E for Cases 1–3 for each day d. For Case 1, C_E was only slightly larger than for Case 2 but considerably smaller than for Case 3. Fig. 8(b) shows the average of the accumulated deviations in T_i^t during the period of every $N_d/4$ days, given by:

$$v_{TC} := \sum_{t=1}^{N_T} \left\{ \max\left(T_i^t - T_{i,\max}^t, 0\right) + \max\left(T_{i,\min}^t - T_i^t, 0\right) \right\} \ , \tag{19}$$

which results from the penalty function of (3). For brevity, (19) is expressed using a linear form, rather than a quadratic form, because v_{TC} can be directly calculated from the optimal profile of T_i^t. In Fig. 8(b), the average of v_{TC} for Case 1 was gradually reduced and became comparable to that for Case 2. As λ_h in (14) increases, v_{TC} can be reduced more rapidly and maintained further lower, although C_E is likely to increase. The case study results verify that the proposed strategy is effective in reducing the HVAC energy cost, while ensuring the thermal comfort.

C. *Comparisons of Operating Schedules of HVAC System*

Fig. 9 represents the 24-h schedules of T_{set}^t and the corresponding variations in P^t and T_i^t for Cases 1, 2, and 3, given the forecasts of C^t and $\mathbf{E}^t$. Specifically, for Case 1, T_{set}^t was scheduled at relatively low levels in the early morning due to the low values of C^t, whereas T_x^t and Q_i^t were maintained high during 7 h ≤ t ≤ 19 h. As C^t began to increase, P^t for Case 1 then became lower than that for Case 3. In other words, the proposed strategy achieved the HVAC load shift from on-peak hours to off-peak hours, leading to the pre-cooling operation and hence the reduction of the HVAC energy cost. Table IV shows that C_E for Case 1 was estimated as \$6.75, which is 28.1% less than \$9.39 for Case 3. Fig. 9(f) shows that in Case 1, T_i^t was still successfully controlled within the acceptable range.

Fig. 10 shows the scheduling results for different profiles of C^t and $\mathbf{E}^t$. Specifically, C^t differed more between the off- and on-peak hours. Fig. 10(a) shows that C^t was negative at t = 3 h and 5 h and increased up to 12.1 ₵/kWh at t = 15 h; note that the y-axis was broken to better display the variation in C^t. Moreover, T_x^t and Q_i^t were estimated to be lower, compared to the cases for Fig. 9(c) and (d), respectively. Therefore, the shift in P^t became larger than for the case of Fig. 9(e), leading to a larger reduction in C_E: i.e., from r_{CR} = 28.1% to 41.7%. In other words, a larger amount of the demand-side flexibility was provided due to the larger difference between C^t for the on- and off-peak hours and the more favorable operating conditions of the HVAC system during the on-peak hours, compared to the case shown in Fig. 9. This confirms that the proposed strategy

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

could successfully reflect the load shifting capabilities of the HVAC system in response to the different profiles of time-varying electricity prices and building thermal conditions.

In Figs. 9 and 10, the optimal schedules for the proposed and ideal strategies (i.e., Cases 1 and 2) were considerably similar, confirming the accuracy of $\mathbf{L}_{1\text{–}3}$ and the convergence of the solution of $\mathbf{S}_3$ to that of $\mathbf{S}_1$, and further to that of the ideal strategy. The small difference arose mainly because the proposed strategy was developed using the actual operating data of the temperature control system, whereas the ideal strategy was achieved using complete information on the system modeling parameters. It was also attributable to the difference between the GD and MILP solvers.

D. Comparisons with Other SL- and RL-based Strategies

The case studies discussed in Sections V-B and V-C were repeated to further evaluate the performance of the proposed SL-based strategy. In particular, as shown in Table V, the proposed strategy was evaluated by comparison with the conventional SL-based strategies, in which a single LSTM was trained offline and online to model the temperature control system. The case with two online-trained LSTMs was also considered, the first of which modeled the thermostat control loop, and the second corresponded to the HVAC system and the building envelope. The comparative study results confirm that the proposed strategy is more effective in improving the building modeling accuracy and temperature control performance, while maintaining the computation time within reasonable limits. The computation time was estimated on a computer with a six-core 4.3-GHz CPU and 32 GB of RAM.

The proposed strategy was also developed using different types of ML models: e.g., one RNN for $\mathbf{L}_1$ and two ARMAXs for $\mathbf{L}_{2,3}$ (Case 4) and two RNNs for $\mathbf{L}_{1,2}$ and one GRU for $\mathbf{L}_3$ (Case 5). The ML models were simpler than the LSTMs. Fig. 11 shows the nRMSEs of $\mathbf{L}_{1\text{–}3}$ for the testing datasets in Cases 4 and 5 over the period from $d = 1$ to N_d. As the online SL continued, the nRMSEs of all $\mathbf{L}_{1\text{–}3}$ were still reduced to low levels in both Cases 4 and 5. After the online SL had finished, the optimal schedules of P^t and the corresponding variations in T_i^t were obtained in Cases 4 and 5 for the profiles of C^t and $\mathbf{E}^t$ shown in Fig. 9. Fig. 12 shows that for Cases 4 and 5, the proposed strategy still achieved the HVAC load shift from on-peak hours to off-peak hours, while leading to small deviations in T_i^t from the acceptable range. This led to a reduction in C_E, compared to Case 3, as shown in Table VI. The case study results confirmed that the proposed strategy can be widely and adaptively applied in real buildings with different temperature control systems and corresponding operating datasets. Moreover, for Cases 4 and 5, the computation times were lower than for Case 1, whereas the HVAC energy costs were higher than for Case 1, revealing the trade-off between the computational burden and the modeling accuracy and scheduling performance.

Furthermore, the proposed strategy was compared with an RL-based strategy using a DDPG algorithm (Case 6) [13], [14]. After initially trained with the historical datasets, the critic and actor networks were further trained for 200 episodes, as in the proposed strategy, each of which was characterized by the profiles of C^t and $\mathbf{E}^t$ for each scheduling day in Case 1. The network weighting coefficients were updated after every episode with the standard deviation of the exploration noise ε_{std} set to 1.2°C. For Case 6, Fig. 12 shows the optimal schedules of P^t and T_i^t, given the profiles of C^t and $\mathbf{E}^t$ shown in Fig. 9, and Table VI lists the corresponding numerical results. For Case 1, C_E and v_{TC} were smaller by 18.6% and 44.0%, respectively, than for Case 6. The computation time for Case 1 was also smaller by 46.4% than for Case 6. The comparative results confirm the advantage of the proposed strategy over the iterative, heuristic strategy. It is worth noting that as in the RL-based strategy, the proposed strategy can directly generate the optimal schedule of HVAC load, when it is integrated with a meta-prediction (MP) method [8]. Briefly, in the MP method, as the datasets of the optimal solutions are collected for various profiles of C^t and $\mathbf{E}^t$, another ANN can be trained to directly generate the optimal schedule. The integration of the proposed

TABLE V. COMPARISONS BETWEEN THE PROPOSED AND CONVENTIONAL SL-BASED STRATEGIES

Profiles of C^t and $\mathbf{E}^t$		Online SL			Offline SL
		3 LSTMs	2 LSTMs	1 LSTM	1 LSTM
Fig. 9	$e(T_i, T_i^{'})$	4.02×10^{-4}	1.53×10^{-3}	3.71×10^{-3}	3.86×10^{-2}
	C_E [\$]	6.75	7.11	8.11	10.6
	r_{CR} [%]	28.1	31.6	22.0	-1.92
	v_{TC} [°C]	1.03	1.20	1.34	0.55
	comp. time [s]	1,618	1,377	1,052	1,073
Fig. 10	$e(T_i, T_i^{'})$	3.97×10^{-4}	1.24×10^{-3}	7.09×10^{-3}	3.82×10^{-2}
	C_E [\$]	4.90	5.77	6.07	8.18
	r_{CR} [%]	41.7	31.4	27.8	2.73
	v_{TC} [°C]	0.18	0.60	0.72	0
	comp. time [s]	1,616	1,442	1,056	1,078

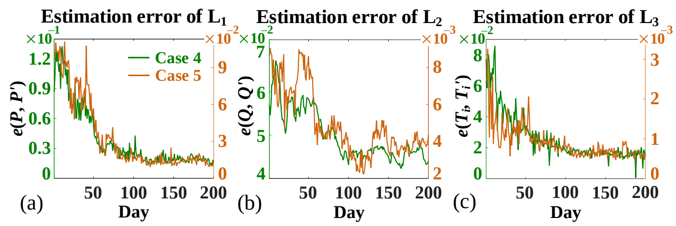

Fig. 11. Variations in the nRMSEs for Cases 4 and 5 during the online SL.

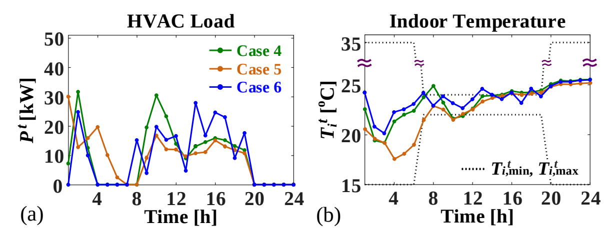

Fig. 12. Comparisons of the 24-h schedules for Cases 4–6: (a) P^t and (b) T_i^t for the profiles of C^t and $\mathbf{E}^t$ shown in Fig. 9.

TABLE VI. COMPARISONS FOR THE PROPOSED STRATEGY WITH DIFFERENT ML MODELS AND THE EXISTING RL-BASED STRATEGY

C^t and $\mathbf{E}^t$ in Fig. 9	Proposed (Case 1)	Proposed (Case 4)	Proposed (Case 5)	RL (Case 6)	Rule-based (Case 3)
$e(T_i, T_i^{'})$	4.02×10^{-4}	2.44×10^{-2}	4.42×10^{-4}	-	-
C_E [\$]	6.75	7.81	6.77	8.29	9.39
r_{CR} [%]	28.1	16.8	27.9	11.7	-
v_{TC} [°C]	1.03	2.78	1.53	1.84	0
comp. time [s]	1,618	918	1,553	3,021	-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

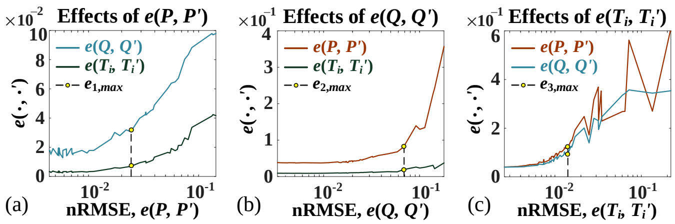

Fig. 13. Effects of the nRMSE of (a) $\mathbf{L}_1$, (b) $\mathbf{L}_2$, and (c) $\mathbf{L}_3$ on the nRMSEs of the other LSTMs.

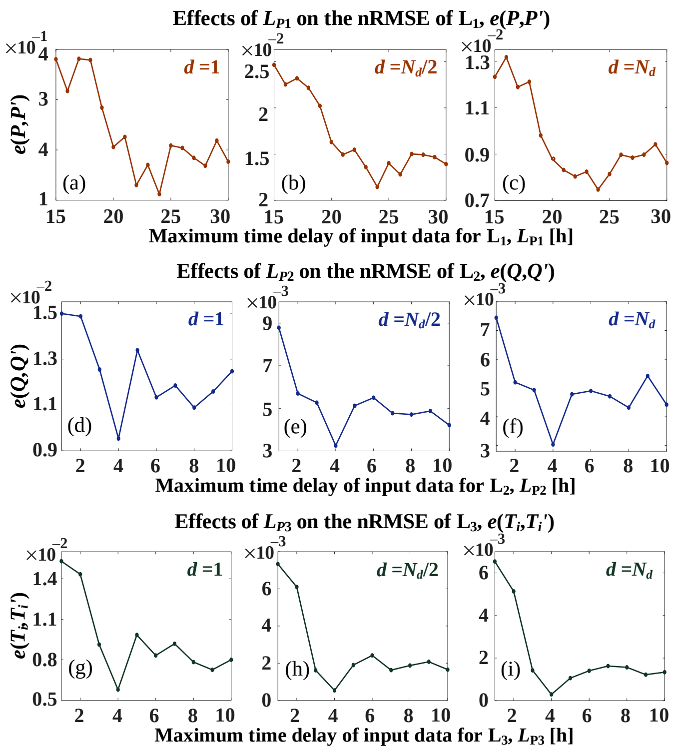

Fig. 14. Effects of the maximum time delays L_{P1}, L_{P2}, and L_{P3} on the nRMSEs of (a)–(c) $\mathbf{L}_1$, (d)–(f) $\mathbf{L}_2$, and (g)–(i) $\mathbf{L}_3$ for d = 1, N_d/2, and N_d.

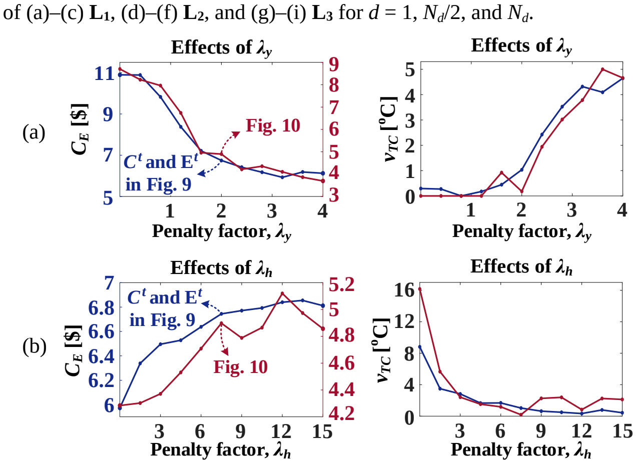

Fig. 15. Effects of the penalty factors (a) λ_y and (b) λ_h on C_E and v_{TC}.

strategy with the MP method has been left for future research.

E. Sensitivity Analyses

For the proposed strategy, the effect of the modeling error of each LSTM was analyzed on the modeling accuracy of the other LSTMs. Fig. 13(a) shows the variations in the nRMSEs of $\mathbf{L}_2$ and $\mathbf{L}_3$ for an increase in the nRMSE of $\mathbf{L}_1$ approximately from 4.40×10^{-3} to 1.45×10^{-1}. When $e(P, P')$ was smaller than about 2.42×10^{-2}, both $e(Q, Q')$ and $e(T_i, T_i')$ marginally increased to low levels. When it became greater than 2.42×10^{-2}, $e(Q, Q')$ and $e(T_i, T_i')$ increased rather rapidly until they were saturated at high levels. This is also the case for the nRMSE variations shown in Fig. 13(b) and (c). Note that in Fig. 13(c), the nonlinearity of the building sub-systems led to the sharp variations in $e(P, P')$ when all the nRMSEs were at high levels. Given the analysis, the permissible error margins of $\mathbf{L}_{1\text{–}3}$ can be specified as 2.42×10^{-2}, 6.79×10^{-2}, and 1.41×10^{-2}, respectively. For all $\mathbf{L}_{1\text{–}3}$, the nRMSEs were smaller than the margins, particularly as the online SL started and continued (see Fig. 7).

In addition, Fig. 14 shows the variations in the nRMSEs of $\mathbf{L}_{1\text{–}3}$ for gradual increases in the maximum time delays of the network input data (i.e., $L_{P1\text{–}3}$ in (6)–(8)). Specifically, Fig. 14(a)–(c) show the variation in $e(P, P')$ with respect to an increase in L_{P1} for d = 1, N_d/2, and N_d, respectively, while L_{P2} and L_{P3} were fixed at 4 h. It can be seen that L_{P1} = 24 h led to the smallest value of $e(P, P')$. When L_{P1} was too small, $\mathbf{L}_1$ could not accurately reflect the thermostat controller operation. Moreover, the operation of the thermostat controller at the current time step was marginally affected by the operations at previous time steps long before the current step. This was also the case for the nRMSE variations in $\mathbf{L}_2$ and $\mathbf{L}_3$ for changes in L_{P2} and L_{P3}, respectively. Fig. 14(d)–(i) show that L_{P2} = 4 h and L_{P3} = 4 h led to the smallest values of $e(Q, Q')$ and $e(T_i, T_i')$, respectively, for d = 1, N_d/2, and N_d. Large values of L_{P2} and L_{P3} did not noticeably improve the accuracy of $\mathbf{L}_2$ and $\mathbf{L}_3$ due to the limited thermal capacity of the test building room.

The case studies, discussed in Section V-C, were also repeated while increasing λ_y and λ_h to 4.0 and 15.0, respectively. Fig. 15(a) represents that an increase in λ_y led to a decrease in C_E and an increase in v_{TC} for both profiles of C^t and $\mathbf{E}^t$ shown in Figs. 9 and 10. When λ_y increased to greater than 2.0, C_E was marginally reduced, whereas v_{TC} was rapidly increased to an inadmissible level. Similarly, Fig. 15(b) shows the case for an increase in λ_h. It can be seen that v_{TC} was slightly reduced when λ_h increased higher than 7.5. Note that the nonlinearity of the sub-systems led to sudden variations in C_E particularly when λ_h varied from 7.5 to 15.0 for the profile of C^t and $\mathbf{E}^t$ shown in Fig. 10. Moreover, for λ_k = 0.5, T_{set}^t and P^t were successfully maintained within the acceptable ranges for all λ_y and λ_h.

VI. Conclusions

This paper proposed a new SL-based strategy for optimal operation of an HVAC system in a commercial building. The system for indoor temperature control was divided into three sub-systems, each of which was modeled using an LSTM. The LSTMs were then interconnected and integrated directly into the optimization problem for temperature set-point scheduling. The optimization problem was reformulated and solved using a deterministic search algorithm within reasonable computation time limits. After optimal scheduling was initiated, the interconnected LSTMs went through the online SL repeatedly, gradually improving the modeling accuracy and the scheduling performance. Case studies were conducted to validate the performance of the proposed strategy in comparison with other strategies using a rule-based temperature set-point, an ideal physics-based building model, and other types of ML-based modeling and scheduling methods. The case study results confirmed that the proposed strategy accurately reflects the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

load-shifting capability of the HVAC system in response to the time-varying electricity prices and building thermal environments, successfully reducing the HVAC energy cost. The results also verified that the proposed strategy effectively mitigates the requirement for historical building data and the risk of unstable operation of the HVAC system and thermal discomfort of occupants in the initial learning period, which is of utmost importance for practical application.

Appendix

A. Consistency Between the Optimal Solutions of $\mathbf{S_2}$ and $\mathbf{S_3}$

The consistency between the solutions of $\mathbf{S_2}$ and $\mathbf{S_3}$ is proved considering a general case [40] as:

$$\text{Minimize } \{ f(\mathbf{x}) : \mathbf{x} \in \mathbf{D} \}, \tag{A1}$$

where f is a nonlinear continuous function on $\mathbb{R}^n$ and $\mathbf{D}$ is a constraint set in $\mathbb{R}^n$. Then, (A1) is reformulated to:

$$\text{Minimize } \{ J(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda \cdot P(\mathbf{x}) \}, \tag{A2}$$

where λ is a positive constant and $P(\cdot)$ is a penalty function on $\mathbb{R}^n$ that satisfies $P(\mathbf{x})$ is continuous; $P(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$; and $P(\mathbf{x}) = 0$ if and only if $\mathbf{x} \in \mathbf{D}$. **Lemma 1** then gives a set of inequalities that follows from the definition of $\mathbf{x}_k = \arg\min_{\mathbf{x}} J(\mathbf{x}, \lambda_k) = \arg\min_{\mathbf{x}} \{f(\mathbf{x}) + \lambda_k \cdot P(\mathbf{x})\}$ and the inequality $\lambda_{k+1} > \lambda_k$.

Lemma 1: $J(\mathbf{x}_k, \lambda_k) \leq J(\mathbf{x}_{k+1}, \lambda_{k+1})$; $P(\mathbf{x}_k) \geq P(\mathbf{x}_{k+1})$; and $f(\mathbf{x}_k) \leq f(\mathbf{x}_{k+1})$

Proof: $J(\mathbf{x}_{k+1}, \lambda_{k+1}) = f(\mathbf{x}_{k+1}) + \lambda_{k+1} \cdot P(\mathbf{x}_{k+1}) \geq f(\mathbf{x}_{k+1}) + \lambda_k \cdot P(\mathbf{x}_{k+1})$

$$\geq f(\mathbf{x}_k) + \lambda_k \cdot P(\mathbf{x}_k) = J(\mathbf{x}_k, \lambda_k), \tag{A3}$$

which proves the first inequality. Moreover, (A3) leads to:

$$f(\mathbf{x}_k) + \lambda_k \cdot P(\mathbf{x}_k) \leq f(\mathbf{x}_{k+1}) + \lambda_k \cdot P(\mathbf{x}_{k+1}) \tag{A4}$$

$$\text{and} \quad f(\mathbf{x}_{k+1}) + \lambda_{k+1} \cdot P(\mathbf{x}_{k+1}) \leq f(\mathbf{x}_k) + \lambda_{k+1} \cdot P(\mathbf{x}_k). \tag{A5}$$

Adding (A4) and (A5) and rearranging the terms yield:

$$(\lambda_{k+1} - \lambda_k) \cdot P(\mathbf{x}_{k+1}) \leq (\lambda_{k+1} - \lambda_k) \cdot P(\mathbf{x}_k), \tag{A6}$$

which proves the second inequality. In conjunction with (A6), the definition of $\mathbf{x}_k$ gives:

$$f(\mathbf{x}_{k+1}) + \lambda_k \cdot P(\mathbf{x}_{k+1}) \geq f(\mathbf{x}_k) + \lambda_k \cdot P(\mathbf{x}_k) \geq f(\mathbf{x}_k) + \lambda_k \cdot P(\mathbf{x}_{k+1}),$$

which proves the third inequality.

Lemma 2: Let $\mathbf{x}^*$ be a solution of (A1). Then, $f(\mathbf{x}^*) \geq J(\mathbf{x}_k, \lambda_k) \geq f(\mathbf{x}_k)$ for each k.

Proof: $f(\mathbf{x}^*) = f(\mathbf{x}^*) + \lambda_k \cdot P(\mathbf{x}^*) \geq f(\mathbf{x}_k) + \lambda_k \cdot P(\mathbf{x}_k) \geq f(\mathbf{x}_k)$.

The two lemmas supports the proof of the theorem on the convergence of the solution of (A2) to that of (A1).

Theorem 1: Let $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_k < \lambda_{k+1} < \cdots \to \infty$. Let $\overline{\mathbf{x}}$ be an arbitrary limit point of $\{\mathbf{x}_k\}_{k=1}^{\infty}$. Then, $\overline{\mathbf{x}}$ solves (A1).

Proof: The limit point is defined as $\overline{\mathbf{x}} = \lim_k \mathbf{x}_k$. Since f is continuous, $\lim_k f(\mathbf{x}_k) = f(\overline{\mathbf{x}})$. Then,

$$J^* := \lim_{k \in \mathbf{K}} J(\mathbf{x}_k, \lambda_k) \leq f(\mathbf{x}^*),$$

$$\Rightarrow J^* = \lim_{k \in \mathbf{K}} f(\mathbf{x}_k) + \lim_{k \in \mathbf{K}} \lambda_k P(\mathbf{x}_k) \leq f(\mathbf{x}^*),$$

$$\Rightarrow J^* = f(\overline{\mathbf{x}}) + \lim_{k \in \mathbf{K}} \lambda_k P(\mathbf{x}_k) \leq f(\mathbf{x}^*),$$

$$\Rightarrow J^* - f(\overline{\mathbf{x}}) = \lim_{k \in \mathbf{K}} \lambda_k P(\mathbf{x}_k) \leq f(\mathbf{x}^*).$$

Since $J^* - f(\overline{\mathbf{x}})$ and $f(\mathbf{x}^*)$ are finite, $\lim_k \lambda_k \cdot P(\mathbf{x}_k)$ is a finite quantity. For $\lambda_k \to \infty$, $P(\mathbf{x}_k)$ converges to zero, verifying $P(\overline{\mathbf{x}}) = 0$.

B. Optimization Problem Constraints for the Ideal Strategy

For the comparative case studies, a physics-based model of the system for the temperature control was implemented as:

$$e^t = T_{set}^t - T_i^{t-1}, \quad \forall t, \tag{B1}$$

$$P_{ref}^t = P_O - k_P e^t - k_I \sum_{\tau=1}^{t} e^{\tau}, \quad \forall t, \tag{B2}$$

$$\sum_{n=1}^{N_S} \delta_n^t = P^t, \quad \forall t, \tag{B3}$$

$$T_i^t = T_{in}^t + \sum_{\tau=1}^{t} \sum_{n=1}^{N_S} F_{n,\tau}^t \delta_n^{\tau}, \quad \forall t, \tag{B4}$$

$$\delta_{n,\max} a_n^t \leq \delta_n^t \leq \delta_{n,\max} a_{n-1}^t, \quad a_0^t = 1, \quad a_{N_S}^t = 0, \quad \forall a_n^t \in \{0,1\}, \forall n, \forall t. \tag{B5}$$

The constraints (B1) and (B2) represent the operation of the PI controller in the thermostat feedback loop. Moreover, (B3)–(B5) correspond to the piecewise linear approximation of the nonlinear variation in T_i^t for a change in P^t [39]. Specifically, in (B3), P^t is divided into N_S linear blocks. In (B4), the variation from T_{in}^t to T_i^t is calculated as the sum of the temperature variations that are led by the incremental HVAC loads assigned in the linear blocks. This is possible because in (B4), $F_{n,\tau}^t$ contains the complete information on the inter-time thermal response of the building to the HVAC system operation. Moreover, (B5) represents the boundaries of the linear blocks to complete the piecewise linear approximation.

References

[1] *Electric Power Monthly with Data for May 2020*, U.S. Energy Inf. Admin., Washington, DC, USA, Jul. 2020. [Online]. Available: https://www.eia.gov/electricity/monthly/current_month/july2020.pdf.

[2] *Commercial Buildings Energy Consumption Survey (CBECS): User's Guide to the 2012 CBECS Public Use Microdata File*, U.S. Energy Inf. Admin., Washington, DC, USA, May 2016. [Online]. Available: https://www.eia.gov/consumption/commercial/data/2012/pdf/user_guide_public_use_may2016.pdf

[3] Z. Afroz *et al.*, "Modeling techniques used in building HVAC control systems: A review," *Renew. Sustain. Energy Rev.*, vol. 83, pp. 64–84, Mar. 2018.

[4] N. Lu, "An evaluation of the HVAC load potential for providing load balancing service," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1263–1270, Sep. 2012.

[5] S. Seyam, "Chapter 4 Types of HVAC systems," in *HVAC System*, London, UK: IntechOpen, 2018, pp. 49–66. [Online]. Available: https://api.intechopen.com/chapter/ pdf-preview/62059

[6] X. Chen, T. Wei, and S. Hu, "Uncertainty-aware household appliance scheduling considering dynamic electricity pricing in smart home," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 932–941, Jun. 2013.

[7] S. Nan *et al*., "Optimal residential community demand response scheduling in smart grid," *Appl. Energy*, vol. 210, pp. 1280–1289, Jan. 2018.

[8] Y. Kim, "A supervised-learning-based strategy for optimal demand response of an HVAC system in a multi-zone office building," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4212–4226, Apr. 2020.

[9] A. Afram *et al.*, "Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems: A state of the art review and case study of a residential HVAC system," *Energy Build.*, vol. 141, pp. 96–113, Apr. 2017.

[10] B. Yuce *et al.*, "An ANN-GA semantic rule-based system to reduce the gap between predicted and actual energy consumption in buildings," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 3, pp. 1351–1363, Jul. 2017.

[11] A. Mathew *et al.*, "Intelligent residential energy management system using deep reinforcement learning," *IEEE Syst. J.*, vol. 214, no. 4, pp. 5362-5372, Dec. 2020.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3051564, IEEE Transactions on Smart Grid

[12] E. Mocanu *et al*., "On-line building energy optimization using deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3698–3708, Jul. 2019.

[13] G. Gao *et al.*, "DeepComfort: Energy-efficient thermal comfort control in buildings via reinforcement learning," *IEEE Internet Things J*., vol. 7, no. 9, pp. 8472–8484, Sep. 2020.

[14] L. Yu *et al*., "Deep reinforcement learning for smart home energy management," *IEEE Internet Things J*., vol. 7, no. 4, pp. 2751–2762, Apr. 2019.

[15] Z. Zhang *et al.*, "Whole building energy model for HVAC optimal control: A practical framework based on deep reinforcement learning," *Energy Build.*, vol. 199, pp. 472-490, 2019.

[16] D. Zhao and Y. Zhu, "MEC—A near-optimal online reinforcement learning algorithm for continuous deterministic systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 2, pp. 346-356, 2014.

[17] G. T. Costanzo *et al.*, "Experimental analysis of data-driven control for a building heating system," *Sustain. Energy Grids Netw.*, vol. 6, pp. 81-90, 2016.

[18] P. Fazenda *et al.*, "Using reinforcement learning to optimize occupant comfort and energy usage in HVAC systems," *J. Ambient Intell. Smart Environ.*, vol. 6, no. 6, pp. 675-690, 2014.

[19] Z. Zou, X. Yu, and S. Ergan, "Towards optimal control of air handling units using deep reinforcement learning and recurrent neural network," *Build. Environ*., vol. 168, pp. 106535, Jan. 2020.

[20] E. F. Camacho and C. Bordons, *Model Predictive Control*. New York, NY, USA: Springer-Verlag, 1999.

[21] Y. Chen *et al.*, "Transfer learning with deep neural networks for model predictive control of HVAC and natural ventilation in smart buildings," *J. Clean. Prod*., vol. 254, pp.1–10, May 2020.

[22] S. Li, S. Ren, and X. Wang, "HVAC room temperature prediction control based on neural network model," in *Proc. IEEE Int. Conf. Meas. Tech. Mech. Aut.*, Hong Kong, China, Jan. 2013, pp. 606–609.

[23] A. Ghahramani *et al.*, "HVAC system energy optimization using an adaptive hybrid metaheuristic," *Energy Build.*, vol. 152, pp. 149–161, Oct. 2017.

[24] W. Chen *et al.*, "Optimal scheduling of demand response events for electric utilities," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 2309-2319, Dec. 2013.

[25] T. Lee, "Optimal spinning reserve for a wind-thermal power system using EIPSO," *IEEE Trans. Power Syst*., vol. 22, no. 4, pp. 1612-1621, 2007.

[26] D. Bertsimas *et al.*, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst*., vol. 28, no. 1, pp. 52-63, Feb. 2013.

[27] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *Int. J. Forecast*., vol. 30, no. 4, pp. 1030–1081, 2014.

[28] S. Rasp *et al.*, "Neural networks for postprocessing ensemble weather forecasts," *Mon. Weather Rev.*, vol. 146, no. 11, pp. 3885-3900, 2018.

[29] Y. Peng *et al.*, "Using machine learning techniques for occupancy-prediction-based cooling control in office buildings," *Appl. Energy*, vol. 211, pp. 1343–1358, Feb. 2018.

[30] C. B. Martinez-Anido *et al*., "The impact of wind power on electricity prices," *Renew. Energy*, vol. 94, pp. 474-487, Jun. 2016.

[31] Z. Y. Wu and T. Walski, "Self-adaptive penalty approach compared with other constraint handling techniques for pipeline optimization," *J. Water Resour. Plan. Manag.*, vol. 131, no. 3, pp. 181-192, May/Jun. 2005.

[32] H. Attouch *et al.*, "Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods," *Math. Program.*, vol. 137, pp. 91–129, 2013.

[33] B. V. Philip *et al.*, "Distributed real-time IoT for autonomous vehicles," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 1131–1140, Feb. 2019.

[34] T. Zakula, "Model predictive control for energy efficient cooling and dehumidification," Ph.D. dissertation, Dept. Archit., Massachusetts Inst. Technol., Cambridge, MA, USA, Jun. 2013.

[35] Y. Kim, L. K. Norford, and J. L. Kirtley, "Modeling and analysis of a variable speed heat pump for frequency regulation through direct load control," *IEEE Trans. Power Syst*., vol. 30, no. 1, pp. 397–408, Jan. 2015.

[36] *Boston, MA Weather History*, Weather Underground. Aug. 2019. [Online]. Available: https://www.wunderground.com/history/daily/us/ma/boston/KBOS/date

[37] K. Ahmed *et al*., "Data for occupancy internal heat gain calculation in main building categories," *Data Brief*, vol. 15, pp. 1030–1034, Dec. 2017.

[38] *Hourly Pricing Program*, ComEd. Aug. 2019. [Online]. Available: https://hourly pricing.comed.com/live-prices/

[39] Y. Kim *et al.*, "Optimal use of thermal energy storage resources in commercial buildings through price-based demand response considering distribution network operation," *Appl. Energy*, vol. 193, pp. 308–324, 2017.

[40] D. Luenberger, *Linear and Nonlinear Programming*, 2nd ed: AddisonWesley, 1989.

Ye-Eun Jang (S'20) received the B.S. and M.S. degree in electrical engineering from Pohang University of Science and Technology, Pohang, South Korea, in 2016 and 2020, respectively, where she is currently pursuing the Ph.D. degree in electrical engineering.
Her research interests include power system reliability, smart grid, and smart buildings.

Young-Jin Kim (Senior Member, IEEE) received the B.S. and M.S. degree in Electrical Engineering from Seoul National University in 2007 and 2010, respectively, and his Ph.D. degree in Electrical Engineering from Massachusetts Institute of Technology in 2015.

He worked for Korea Electric Power Corporation as a power transmission and distribution system engineer from 2007 to 2011. He was also a visiting scholar at the Catalonia Institute for Energy Research in 2014, and a postdoctoral researcher for Center for Energy, Environmental, and Economic Systems Analysis at the Energy Systems Division at Argonne National Laboratory from 2015 to 2016. He joined the faculty of the Pohang University of Science and Technology (POSTECH), where he is currently an Associate Professor in the Department of Electrical Engineering. His research fields of interest include distributed generators, renewable energy resources, and smart buildings.

João P. S. Catalão (Senior Member, IEEE) received the M.Sc. degree from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 2003, and the Ph.D. degree and Habilitation for Full Professor ("Agregação") from the University of Beira Interior (UBI), Covilha, Portugal, in 2007 and 2013, respectively.

Currently, he is a Professor at the Faculty of Engineering of the University of Porto (FEUP), Porto, Portugal, and Research Coordinator at INESC TEC. He was also appointed as Visiting Professor by North China Electric Power University (NCEPU), Beijing, China. His research interests include power system operations and planning, distributed renewable generation, power system economics and electricity markets, demand response and smart grid.