Dynamic Programming based Approaches to Optimal Rooftop Unit Coordination[1]

VAMSI PUTTA[1], DONGHUN KIM[2], JIE CAI[2], JIANGHAI HU[1], JAMES E BRAUN[2]
[1]*School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA*
[2]*School of Mechanical Engineering, Purdue University, West Lafayette, IN, USA*

While buildings served by multiple rooftop units (RTUs) are widely prevalent in retail and food industries, optimal coordination among RTUs has received little attention. Currently, most buildings utilize plain feedback or heuristic strategies to determine RTU operations. In this paper, we present a switched dynamic programming based approach for optimal scheduling of rooftop unit operations. Using a mid-size restaurant served by multiple RTUs as an example, a switched affine dynamic system model is first presented. The RTU coordination control problem is then formulated as a switched optimal control problem with nontrivial switching cost. A dynamic programming based method is proposed for the solution of the switched optimal control problem, which has reasonable computational complexity and can be implemented either as a one-shot solution or as part of a model predictive control algorithm. In order to mitigate the growing computational complexity of the proposed algorithm, we extend the approach by proposing a distributed extension of the model predictive control formulation. Finally we conduct simulation-based evaluation of the various formulations and point out the salient features of the formulation.

INTRODUCTION

Rooftop Units (RTUs) are packaged air handling units that regulate the temperature and circulate air as part of a Heating, Ventilation and Air-Conditioning (HVAC) system. Rooftop Units are among the most prevalent HVAC systems in commercial buildings in the United States especially in the retail and food sectors. In a majority of such buildings, each RTU's operation is controlled by cycling it on and off to attain the temperature setpoint in a particular zone of the building. Heuristic control strategies of individual RTUs without considering mutual coordination can lead to wasteful scenarios where different RTUs compete (simultaneously heating and cooling) to maintain their respective temperature setpoints, resulting in higher energy expenditure.

---

[1] A preliminary version of this paper was submitted to the 2015 American Control Conference.

A building with multiple RTU units operating in on/off fashion can be modeled as a switched affine system (SAS). SASs are a class of hybrid dynamical systems comprising multiple affine dynamical subsystems whose control is specified by a (discrete) switching sequence in addition to a continuous control input. A building with multiple RTUs operates in a number of distinct modes, one corresponding to an ON/OFF configuration of the RTUs. In each mode, the building thermal dynamics can be modeled by affine dynamics obtained from a RC network approximation. The mode sequence over any given time horizon specifies the operating schedule of the RTUs. An optimal control problem can then be formulated to find the operation schedule (and the continuous control, if applicable) that minimizes the energy cost of the RTUs while maintaining comfort for building occupants. Another cost that needs to be considered is the loss in lifetime of the RTUs due to their frequent turning on and off, which can be modeled as switching cost in the SAS optimal control problem formulation.

Dynamic programming has been one of the primary methods for solving optimal control problems, including those for switched and hybrid systems [1]. For a switched optimal control problem such as the RTU coordination problem under consideration, a key challenge lies in its combinatorial nature: the number of operation schedules that need to be considered increases exponentially with the prediction time horizon. Thus, straightforward application of dynamic programming results in exponential complexity growth, making the problem intractable for long time horizons [2]. Mitigations in the form of relaxation [1] and tree pruning [3] have been proposed to reduce the growth of complexity. An application of relaxed dynamic programming with stability analysis to the Model-based Predictive Control (MPC) of switched linear systems was presented in [4]. Nevertheless, practical applications of these methods remain few to date.

In this paper, a model-based predictive control algorithm is proposed for the coordination of multiple RTUs having different efficiencies in a multi-zone building. MPC algorithms have increasingly become attractive options in building control [5, 6] due to their ability to utilize real-time weather and occupant information to minimize energy consumption. In the proposed solution, the RTU coordination problem over a given time horizon is formulated as a discrete-time switched affine quadratic regulation (SAQR) problem with mode-dependent switching costs and solved by the dynamic programming method with

complexity reduction techniques. Simulation results show that the proposed approach can lead to reduced energy expenditure of RTUs through a better coordination among them.

This paper is organized as follows. In the following section we describe the building that serves to motivate this study. We utilize an existing high fidelity model of this building to obtain a model that is amenable to control design. From the knowledge of the building capabilities, we formulate the optimal control problem as a switched linear quadratic regulator problem. We then introduce the proposed dynamic programming approach and variations to solve the optimization problem. In particular, we describe a distributed method of optimization in order to mitigate the complexity growth rate. Controller implementations of the proposed methods are then discussed with a simulation-based evaluation. Finally, we draw conclusions about the strengths and weaknesses of the proposed approach.

CASE STUDY BUILDING

We now describe the building that motivates this study and doubles as a test case for the proposed approach. The building under consideration is a medium-size restaurant that is served by 4 RTUs of varying capacity and efficiency. Fig. 1 depicts the internal layout of the restaurant, with the numbers indicating the RTUs that serve the dining area. RTU 1 feeding the main dining area is the largest and the most efficient unit while RTU 2, 3 and 4 are identical smaller units serving the smaller zones. The capacity and efficiency (COP) of these RTUs are summarized in Table 1.

In the existing configuration of the restaurant on site, air temperature is sensed at each of the four thermostat locations, and the corresponding RTU cycles its compressor ON or OFF based on the deviation of the measured temperature from a user specified setpoint without coordinating with other RTUs. All the RTU fans are on continuously. In this study, it is assumed that each RTU provides cool air at a given supply air temperature when they are ON and air at ambient temperature when OFF. This assumption significantly simplifies the effect of the RTU operation sequence on the building thermal dynamics and vice versa. Each RTU is also assumed to incur a fixed cost of operation that is proportional to its energy consumption when ON and zero cost when OFF, thus ignoring the time variance of the efficiencies. While these assumptions are strictly speaking not true, the error incurred is small for the relatively small time horizon control which is the focus of this paper.

A high-fidelity Computational Fluid Dynamics (CFD) coupled model of the restaurant's building envelope was developed and validated based on the approach presented in [7, 8]. However, the resulting model had a high state dimension for controller design purposes. Hence assuming this model to be the ground truth, a four-dimensional linear system model (inverse model) was developed using the subspace identification (n4sid) method in MATLAB. The training data was obtained by simulating a typical feedback control strategy for a one-week period designed to keep the zone temperatures sensed at the thermostats (available as outputs from the high-fidelity model) at a comfortable level between 21° C and 26° C. The TMY3 weather profile for June 2014 was used to model the ambient conditions, while a typical occupancy profile was assumed in the training process making the set of inputs representative of data available on-site. The resulting inverse model was validated by using one week of the July 2014 weather profile. When compared to the high-fidelity CFD coupled model, the trained linear model had an RMSE of 3°C across all zones over the week. When restricted to the occupancy period the RMSE decreased to 1.1^{o}C which implies that occupant comfort would not be affected too adversely. Figure 2 depicts the dining area temperature measurement during the validation period. In the next section, we describe how the obtained inverse model can be formulated as a switched affine system.

SWITCHED AFFINE MODEL

We begin by describing a general state space model of the building envelope. Distinguishing between controllable inputs and exogenous inputs, the building thermal dynamics can be modeled as

$$x(t+1) = Ax(t) + Bu_{\sigma(t)} + Fw(t) \tag{1}$$

$$y(t) = Cx(t), \qquad t = 0,1,\ldots.$$

Here, the state $x(t) \in \mathbb{R}^n$ is an n-dimensional vector consisting of the temperatures of suitable wall and air nodes in the CFD model. Since the inverse model is obtained through training, the actual components of $x(t)$ will be in general combinations of the node temperatures. The operation schedule $\{\sigma(t)\}, t = 0,1,\ldots$ is a sequence of variables with a finite set of possible values (denoted Σ), one corresponding to each possible combined ON/OFF status of all the RTUs. For the case study building, we model the four RTUs by Σ consisting of a total of 16 possible values (or modes). The exogenous input $w(t)$ models external heat gains from solar radiation, as well as internal gains from lighting and occupants. The controllable input $u_{\sigma(t)}$ is a

vector of zeros and ones with each element corresponding to the ON/OFF status of an RTU for the mode $\sigma(t) \in \Sigma$ at time t. The output variable $y(t) \in \mathbb{R}^p$ contains the air temperature measurements at the thermostat locations. For the case study, this is a four dimensional vector $(p = 4)$. $A, B, F,$ and C are constant matrices of proper dimensions obtained via training in the case of the inverse model.

Considering that the controllable inputs $u_{\sigma(t)}$ can only take binary values, we can simplify the above model by combining the contribution of all the RTUs to the envelope dynamics as follows.

$$x(t + 1) = Ax(t) + b_{\sigma(t)} + Fw(t), \tag{2}$$

$$y(t) = Cx(t), t = 0,1, \ldots.$$

Here, $b_{\sigma(t)} = Bu_{\sigma(t)}$ whose value depends only on the current RTU ON/OFF configuration $\sigma(t)$. Due to the presence of the $Fw(t)$ term and the mode sequence $\{\sigma(t)\}$ as the control input, this is an instance of the switched affine systems (SASs).

To facilitate the study of the ensuing temperature tracking problem with possibly time-varying setpoint offsets, the above SAS model can be reduced to a simpler Switched Linear System (SLS) model through a standard homogenization procedure, if the exogenous input w(t) is assumed to be known or predictable within the given time horizon. To this end, assume that $x_{set}(t)$ is a setpoint (reference) trajectory of the state and $y_{set}(t) = Cx_{set}(t)$ the corresponding setpoint output trajectory. Define the augmented and offset state and the offset output as

$$\hat{x}(t) := \begin{bmatrix} x(t) - x_{set}(t) \\ 1 \end{bmatrix}, \qquad \hat{y}(t) := y(t) - y_{set}(t) \tag{3}$$

respectively. Then (3) can be written as

$$\hat{x}(t + 1) = \hat{A}_{\sigma(t)}\hat{x}(t), \qquad \hat{y}(t) = \hat{C}\hat{x}(t), \tag{4}$$

where $\hat{A}_{\sigma(t)} = \begin{bmatrix} A & b_{\sigma(t)} + Fw(t) + Ax_{set}(t) - x_{set}(t + 1) \\ 0 & 1 \end{bmatrix}$ and $\hat{C} = [C \;\; 0]$.

Note that the system in (4) is now a homogeneous switched linear system. Under our assumption that the exogenous inputs are completely predictable, $\hat{A}_{\sigma(t)}$ is known or predictable with values dependent on the mode sequence σ(t) within the given time horizon.

PROBLEM FORMULATION

For the building system described in the previous section, the optimal RTU coordination problem to be formulated in this section is an optimal control problem of the SAS (2) or the SLS (4) whose objective is to minimize the operational cost while maintaining occupant comfort during a given time horizon. For simplicity, we assume in this section that the current time is t = 0 and the given time horizon is t =0,1,...,k. We first discuss three factors that contribute to the cost function over this time period.

a) Energy Cost: For a building with r RTUs, each mode σ ∈ Σ can be represented by a binary string $\sigma_1\ \sigma_2\ \cdots\ \sigma_r$ where $\sigma_j\ =\ 0$ and 1 represents the OFF and ON status of RTU j ∈ {1,...,r}, respectively. Assume that the power consumed by RTU j is p_j when it is ON and 0 when it is off. Then the total energy consumed by the RTUs during the period t = 0,1,...,k is given by

$$J_e\ (k) = \sum_{t=0}^{k} \lambda(t) p_{\sigma(t)} \Delta$$

where $\lambda(t)$ is the (predicted) utility price at time t, Δ is the sampling time, and $p_{\sigma(t)}$ is the total power consumed by the RTUs in mode σ(t):

$$p_{\sigma(t)} = \sum_{j=1}^{r} \sigma_j(t) p_j$$

b) Comfort Penalty: Another factor that needs to be considered in the cost function is the occupant comfort. There are various existing metrics to measure occupant (dis)comfort, e.g., Percentage of People Dissatisfied (PPD) and Predicted Mean Vote (PMV) [9, 10, 11]. In this paper, a simplified approach is adopted: we assume that a reference trajectory of the output (i.e., setpoint temperatures at the thermostat locations) $y_{set}\,(t)\ =\ Cx_{set}\,(t)$ is given over the time horizon $t\ \in\ \{0,\ldots,k\}$; and deviation from it at any time t will incur a penalty $[y(t) -\ y_{set}(t)]^T H(t)[y(t)\ - y_{set}\,(t)]$ for some positive semidefinite matrix $H(t)$. Thus, the total comfort penalty cost over the time period is

$$J_c(k)\ = \sum_{t=0}^{k} [y(t) -\ y_{set}(t)]^T H(t)[y(t)\ - y_{set}\,(t)]$$

c) Switching Cost: In practice, an RTU has a finite lifespan, which may be shortened by frequently turning it on and off. In addition, turning several RTUs on at the same time may lead to large spikes in power demand, which could result in excessive demand charges in the utility bill. Thus, the optimal control problem formulation needs to take into account the cost associated with the RTUs switching

status at each time step. Specifically, for each RTU $j \in \{1, \ldots, r\}$, denote by c_{01}^{j} (resp. c_{10}^{j}) the cost associated with it switching from OFF to ON (resp. from ON to OFF) at each time step.

Let σ(t) and σ(t + 1) be two consecutive modes. Then the cost of switching from σ(t) to σ(t + 1) is

$$c_{\sigma(t),\sigma(t+1)} := \sum_{t=0}^{k} c^{j}{}_{\sigma(t),\sigma(t+1)}$$

Thus, the total switching cost during the time period $t \in \{0, \ldots, k\}$ is

$$J_s(k) := \sum_{t=0}^{k} c_{\sigma(t),\sigma(t+1)}$$

where for notational simplicity later on we have assumed that σ(k+1) = σ(k), i.e., there is no switching after the last time step.

Optimal Control Problem Formulation

With the costs defined above, the optimal RTU coordination problem can be formulated as the following optimal control problem:

$$\text{Minimize } J(k) = J_e(k) + J_c(k) + J_s(k)$$

subject to system dynamics (2) for $t = 0,1, \ldots, k$.

Note that the optimal control to be solved is the RTU operation schedule σ(t) over the horizon t ∈ {0,...,k} with the understanding that σ(k + 1) = σ(k). By using the augmented and offset state and output defined in (3), the SAS dynamics (2) are simplified to the SLS dynamics (4); the cost function J(k) is also reduced to a quadratic one,

$J(k) = \sum_{t=0}^{k} \hat{x}(t)^T \widehat{H}(t) \hat{x}(t)$ where $\widehat{H}(t)$ is defined as

$$\widehat{H}(t) = \begin{bmatrix} C^T H(t) C & 0 \\ 0 & \lambda(t) p_{\sigma(t)} \Delta + c_{\sigma(t),\sigma(t+1)} \end{bmatrix}$$

As a result, the optimal RTU coordination problem can be equivalently formulated as

$$\text{Minimize } J(k) = \sum_{t=0}^{k} \hat{x}(t)^T \widehat{H}(t) \hat{x}(t) \tag{5}$$

subject to $\hat{x}(t + 1) = \hat{A}_{\sigma(t)} \hat{x}(t), \quad \hat{y}(t) = \hat{C} \hat{x}(t)$

Note that problem defined by equation (5) can be considered as a generalized version of the switched linear quadratic regulation (SLQR) problem studied in [2] in that the quadratic matrix $\widehat{H}(t)$ for the cost at time t depends not only on the current mode $\sigma(t)$ but also on the next mode $\sigma(t\ +\ 1)$.

DYNAMIC PROGRAMMING SOLUTION

The optimal control problem (5) formulated in the previous section can be solved using a dynamic programming method to be presented in this section, which is a (slight) generalization of the algorithm proposed in [3]. Denote by $V_s\ (\hat{x})$ the value function (cost-to-go) of the problem (5) over the time horizon $\{s, s\ +\ 1, \ldots, k\}$:

$$V_s\ (\hat{x}) := \min_{\sigma(t),\ \ s\le t\le k} \left\{ \sum_{t=s}^{k} \hat{x}(t)^T \widehat{H}(t)\hat{x}(t)\ \ \mid \hat{x}(s) = \hat{x} \right\}$$

for $\hat{x} \in \mathbb{R}^{n+1}$ and $s\ \in\ \{0,1, \ldots, k+1\}$. Then $V_s\ (\hat{x})$ satisfies the Bellman equation

$$V_s\ (\hat{x}) = \min_{\sigma(s)\in\Sigma} \left\{ \hat{x}(s)^T \widehat{H}(s)\hat{x}(s)\ + V_{s+1}\left(\hat{A}_{\sigma(s)}\ \hat{x}\right)\ \right\} \tag{6}$$

for all $\hat{x} \in \mathbb{R}^{(n+1)\times(n+1)}$and s ∈ {0,1,...,k}, with zero terminal cost $V_{k+1}\ (\cdot) \equiv 0$ due to our assumption that σ(k+1) = σ(k). It should be pointed out that, in the Bellman equation (6), the matrix $\widehat{H}(s)$ also depends on $\sigma^*(s\ +\ 1)$, which is the optimal mode when starting from the state $\hat{x}(s\ +\ 1) := A_{\sigma(s)}\ \hat{x}$ at the next time step $s\ +\ 1$:

$$\sigma^*(s\ +\ 1) = \underset{\sigma(s)\in\Sigma}{\arg\min}\ \ \left\{ \hat{x}(s\ +\ 1)^T \widehat{H}(s+1)\hat{x}(s+1) \right. \tag{7}$$

$$\left. + V_{s+2}\ \left(\hat{A}_{\sigma}(s+1)\hat{x}(s+1)\right)\right\}$$

for s ∈ {0,1,...,k −1}. Note that $\sigma^*(s+1)$ itself depends on σ(s). When $s\ =\ k,\ \sigma^*(k\ +\ 1)\ =\ \sigma(k)$. To sum up, the special structure of the $\widehat{H}(t)$ matrix in the optimal RTU coordination problem renders the iteration (6) a less straightforward process than the conventional Bellman iterations in that the control decision σ(s) affects the running cost in a complicated way through the dependency of $\widehat{H}(s)$ on $\sigma^*\ (s\ +\ 1)$ and hence ultimately on $A_{\sigma\ (s)}\ \hat{x}$.

Carrying out the iteration (6) for $s\ =\ k, k\ -1, \ldots, 0$ with the terminal conditions $V_{k+1}\ (\cdot)\ \equiv\ 0$ and $\sigma^*(k\ +\ 1)\ =\ \sigma(k)$, we can obtain all the value functions $V_s\ (\cdot)$, in particular, $V_0\ (\cdot)$. The optimal cost for

problem (5) is then $V_0\left(\hat{x}(0)\right)$. The optimal mode sequence $\{\sigma^*(t)\}\ t = 0, \ldots, k$ resulting in the optimal cost can be recovered by a forward iteration, yielding the optimal operation schedule for all the RTUs.

REPRESENTATION OF VALUE FUNCTIONS

For SLQR problems without switching cost, it was shown in [2] that their value functions are the pointwise minimum of finite families of quadratic functions, hence piecewise quadratic in themselves. For the RTU coordination problem (5) with switching cost, such a property still holds.

Proposition 1: For each $s \in \{0,1,\ldots,k\}$, the value function $V_s\,(\cdot)$ can be represented as

$$V_s(\hat{x}) = \min_{Q \in Q_s} \hat{x}^T Q \hat{x}. \tag{8}$$

for some finite set Q_s of positive semi-definite matrices in $\mathbb{R}^{(n+1)\times(n+1)}$.

Proof: The complete proof can be found in [2]. Here we only provide a sketch of the proof.

For a fixed operation schedule, the dynamics (4) becomes a linear time-varying system, and all the quadratic matrices $\widehat{H}(t)$ are fixed, resulting in quadratic costs-to-go. For the switched optimal control problem (5), each value function is the minimal of such costs-to-go over all possible operation schedules. This proves the desired result.

Using the representation (8), the value function iterations (6) then reduce to iterative procedures for obtaining the sets $Q_k\,, Q_{k-1}, \ldots, Q_0$ as follows

$$Q_s \;=\; \rho_\Sigma^s\left(Q\,(s+1\,)\right), \qquad s \;=\; k, k\,-\,1, \ldots, 0.$$

with the terminal condition $Q_{k+1} = \{0\}$. Here, ρ_Σ^s is the switched Riccati mapping at time s defined by

$$\rho_\Sigma^s\;(Q_{s+1}) \coloneqq \left\{\rho^s{}_{\sigma,\sigma'}(Q) \mid \sigma, \sigma' \in \Sigma, Q \in Q_{s+1}\right\} \tag{9}$$

where the Ricatti mapping $\rho^s{}_{\sigma,\sigma'}$ is defined for $\sigma, \sigma' \in \rho$ as

$$\rho^s{}_{\sigma,\sigma'}(Q) = \;\hat{A}_\sigma(s)^T Q \hat{A}_\sigma(s) + \widehat{H}(s) \in \mathbb{R}^{(n+1)\times(n+1)} \tag{10}$$

for any $Q \in \mathbb{R}^{(n+1)\times(n+1)}$. Note that in (10), $\hat{A}_\sigma(s)$ depends only on σ while $\widehat{H}(s)$ depends on the mode pair (σ, σ'). In the case of zero switching cost, $\widehat{H}(s)$ depends only on the current mode $\sigma(s)$; hence σ' can be dropped from both of the definitions (9) and (10).

The switched Riccati mapping (9) provides a way of computing the sets Q_s , hence the value functions $V_s\,(\cdot)$, iteratively. However, the complexity of the representation (8) as measured by the cardinality of the

set Q_s grows exponentially with the iterations. Several complexity reduction techniques were proposed in [3], including pruning at each step redundant matrices that do not contribute to the minimum in (8), and relaxed dynamical programming methods that remove almost redundant matrices at the expense of accuracy. These techniques after slight modification can be applied to reduce the computational complexity of the iteration (9). For more details, the interested readers can refer to [3]. Here we propose a distributed optimization routine that trades off optimality to achieve complexity reduction.

DISTRIBUTED OPTIMIZATION

By decomposing the optimization problem into independent sub-problems, one can achieve complexity reduction especially if the sub-problems can be solved in parallel. However, given the coupled nature of cost functions, such decomposition might not be evident. Distributed optimization methods aim to overcome this hurdle by approximating the original coupled cost function as a sum of decoupled sub costs each of which can be minimized in parallel. The approximations are then updated to reflect the coupling more accurately and the optimization repeated. Distributed approaches to MPC are receiving increased attention recently [12, 13, 14, 15]. Apart from complexity reduction, distributed approaches can also facilitate plug-and-play and easily reconfigurable controls.

In the current study, the coupling in the cost function occurs via the building thermal dynamics. While the RTU power consumptions are independent, the zone temperatures react to all the RTU inputs. Hence it is not trivial to assign accurate contributions of each RTU to the total cost function. However, one observes that if all but one RTU are fixed at optimal operation, then the optimization problem (5) incurs a reduction in the degrees of freedom and transforms to optimizing a single RTU switching. If we denote the switching sequence of RTU i with σ_i and its optimal counterpart using the superscripted σ_i^*, we can write

$$V_s = \min_{\sigma_i(s)\in\{0,1\}} \left\{ \hat{x}(t)^T \hat{H}(t)\hat{x}(t) \ + V_{s+1}\left(\hat{A}_{[\sigma_i(s) \oplus \sigma^*_{r/i}(s)]}\, \hat{x}\right) \right\}$$

Here the total switching sequence $\sigma(s)$is written as the direct sum $\left[\sigma_i(s) \oplus \sigma^*_{r/i}(s)\right]$ with all RTUs but RTU i optimized. Note that V_s now has only two degrees of freedom corresponding to $\sigma_i(s) = 0/1$. Hence applying the switched Ricatti mapping would only cause the family $Q_s$to grow as a power of 2 rather than 2^r.

Since the optimal $\sigma_{r/i}^{*}$ is not known apriori, one can choose a reasonable initial guess $\sigma_{r/i}^{init}$ and optimize to obtain $\sigma_i^{*}(\sigma_{r/i}^{init})$. Note the dependence of the σ_i^{*} on the initial guess. Also $\sigma_i^{*}\left(\sigma_{r/i}^{*}\right) = \sigma_i^{*}$. Hence we aim to get $\sigma_{r/i}^{init}$ to converge to σ_j^{*}. Also note that, once an initial guess for the RTU operations is chosen, we can obtain individual copies of the value function corresponding to each RTU as follows.

$$V_s^i = \min_{\sigma_i(s)\in\{0,1\}} \left\{ \hat{x}(t)^T \hat{H}(t)\hat{x}(t) + V_{s+1}\left(\hat{A}_{[\sigma_i(s) \oplus \sigma_{r/i}^{init}(s)]}\,\hat{x}\right) \right\}, \qquad s = 0,1,\dots,k$$

As defined, these individual value functions denoted by V_s^i form an over approximation and converge to the actual value function as $\sigma_{r/i}^{init}$ converges to $\sigma_{r/i}^{*}$. The individual value functions can be optimized simultaneously to obtain $\sigma_i^{*}(\sigma_{r/i}^{init})$ which can be used to update the initial guess $\sigma_{r/j}^{init}$ at RTU j. Repeating the process by updating the initial guess at each RTU can lead to the individual value functions converging to the optimal value function. This idea is summarized in Algorithm 1.

Remark: Convergence of the individual value functions to the optimal value function is not guaranteed due to the coupling present in the cost function. However, with a reasonable initial guess, it is possible to consider only those valid iterations that present an improvement in the value function from the previous iteration. This leads to a non-increasing sequence of value functions with each valid iteration.

Algorithm 1

1. Choose an initial guess σ_r^{init} for the operation schedule of all RTUs
2. For $j = 1,2,\dots,j_{max}$
    a. For $i = 1,\dots,r$
        i. At RTU i, optimize V_0^i to obtain $\sigma_s^{*}\left(\sigma_{r/i}^{init}\right), s = 0,1,\dots,k$
    b. End For
    c. Update σ_r^{init} with $\sigma_i^{*}(\sigma_{r/i}^{init})$
3. If σ_i^{*} has converged for all i, RETURN
4. End For

Algorithm 1 presents the distributed dynamic programming approach to MPC. Since each V_0^i has only one degree of freedom, the rate of growth of matrices is exponentially slower than the centralized approach. This reduction in computational complexity comes with suboptimality as V_0^i are over approximations and in

fact might not converge to V_0. However, the distributed solution does provide a decrease in the value function compared to the initial guess and with a sufficient number of iterations can achieve near optimality.

In the next section we discuss the simulation results obtained from implementing the proposed dynamic programming approach to the high fidelity model.

SIMULATION RESULTS

In this section, the performance of the proposed algorithms will be tested on the high fidelity restaurant model. The savings in power cost and temperature regulation behaviors for the different implementations of the proposed algorithm will be compared.

All MPC simulations discussed here were simulated using the 4-dimensional black-box model of the restaurant to generate control input, which was fed to the high-fidelity model. Since there was no one-to-one correspondence between the states of the high-fidelity model and black-box model, it was not possible to use feedback from the high-fidelity model in the optimal control design. Rather the black-box model was used to propagate the state trajectories for look-ahead. This corresponds to a situation where optimal control must be implemented without any sensors or observers. To reduce the mismatch between initial states of the two models, a week of warmup was provided for all the simulations.

A look-ahead horizon of 6 (12) hours (k=36 (72) sampling instants with a sampling time of 10 minutes) is used for the centralized and distributed MPC, respectively, with perfect prediction of the exogenous inputs. The longer lookahead of the distributed MPC emphasizes its potential in reduction of computational complexity. The comfort penalty quadratic matrix $H(\cdot)$ was chosen to be $7I_4$, where I_4 denotes the 4-dimensional identity matrix. The temperature setpoint y_{set} was chosen to be 23° C for all zones. Each RTU was assigned a switching cost corresponding to 15% of its power consumption. This reflects a higher penalty in cycling the largest RTU on and off, thus implicitly modeling a demand charge based scenario where consumers are billed on the maximum power usage.

MPC approaches include updating the state and exogenous input estimates at each time step and can adapt quickly to changes, which is an important feature in applications with uncertain information. Alternately if accurate prediction of the exogenous inputs is available over a long time horizon, the optimal

control problem can be solved only once, with the resulting optimal RTU operating schedule implemented during the whole look-ahead horizon (rather than just its first step as in the MPC case). In this case, the problem becomes a finite-horizon switched discrete-time LQR problem that needs to be solved only infrequently and its solution can be carried out offline at non-critical times such as at night when the optimal control is straightforward, leading to real-time implementation. We term such a control approach as a one-shot MPC.

Simulations were performed to obtain the power consumptions of the three control implementations during a one-week period. Performance in maintaining comfort is quantified in terms of the RMSE deviation of the zone temperatures from the setpoints during the occupancy hours, averaged across the four zones. A simple feedback based algorithm, where the RTUs are cycled based on the temperature measurement at the corresponding thermostats was also simulated to serve as the baseline. This baseline simulation was implemented completely on the high-fidelity model to accurately represent the nominal building energy costs. A deadband of $0.5^{o}C$ was included in the control logic to prevent chattering. Tree pruning with a relaxation parameter of 0.01 was used to keep the computation feasible in real time for the dynamic programming algorithm. All controllers were designed to operate from 6:00 am till midnight with the temperature being allowed to float during the remaining times.

Figures 3 and 4 depict the performance of the controllers in terms of maintaining comfort and the corresponding RTU operations. As can be seen the controllers are able to maintain reasonable thermal comfort. The one-shot and the MPC controllers offer some precooling in anticipation of the upcoming internal gains at the start of the working hours. The conventional bang-bang controller slightly outperforms the other controllers in temperature regulation by virtue of focusing on temperature tracking alone. During the occupancy period, the centralized controllers are observed to exhibit more cycling compared to the distributed controllers with a longer horizon. This is due to the choice of the switching cost. Longer look-ahead controllers amortize the switching costs over a longer horizon by saving power costs. The look-ahead based controllers turn on the more efficient RTU #1 earlier and run it longer which is reflected in the power savings. The conventional bang-bang controller exhibits longer run times across all RTUs. The included deadband helps minimizing short cycling and can be adjusted as necessary.

Table 2 summarizes the mean energy consumption per day and the temperature regulation across all the zones. The one-shot controller offers the most savings (3.7%) compared to the baseline bang-bang control. The centralized 6-hour look-ahead MPC comes close to matching the savings (3.6%) while the distributed solver with longer horizon performs on par with bang-bang control. All three controllers perform reasonably well in thermal regulation with the average deviation being less than 2° C though the bang-bang controller provides tighter thermal regulation. It must be noted that, this simulation tests the performance of the controller without any feedback from the plant and the performance can be improved with a state observer or sensors. The relatively high internal gains required the RTUs to typically remain in the ON state for most of the occupancy hours, leading to reduced potential for savings.

The discrepancy between the savings incurred due to the centralized and distributed solvers was further investigated. By simulating the distributed controller using the black box model in lieu of the high fidelity model, the savings increased by 2%, suggesting that the long horizons involved in the distributed control amplified the mismatch between the simulation and control models. This mismatch can typically be avoided by using a suitable observer for the states. Furthermore, allowing the information exchange iterations to continue till convergence yielded an extra 0.3% savings over a week period showing that premature truncation lost some benefits of the distributed control. Distributed MPC can still make realtime computations feasible especially in large scale problems as observed from the solution complexity reported next.

SOLUTION COMPLEXITY

Due to the combinatorial nature of the optimization problem, the number of matrices required to represent the value function grows exponentially with the time horizon. For the current simulation, we mitigate this complexity growth by pruning and relaxing the value function computation [2]. Using a relaxation threshold of 0.01 yields approximately 850 matrices in the representation of the value function after 36 iterations (6 hour look-ahead with a sampling time of 10 minutes) compared to the theoretical upper bound of 16^{36} matrices. On a 2.4Ghz Intel Core i3 processor based PC, the maximum computation time to solve the MPC problem at any sampling instant was 8.8 minutes which points to real time feasibility considering the 10 minute sampling interval. For distributed computation, pruning was

unnecessary. However, to maintain real time computational capability, information exchange was limited to a maximum of 4 iterations rather than waiting for convergence amongst the local controllers. This still yielded a solution for the 12 hour look-ahead MPC in the same time as the centralized 6 hour look-ahead MPC.

CONCLUSIONS

Applying concepts from switched linear systems to RTU control enables us to formulate and solve the problem for sufficiently long look-ahead horizons to achieve savings. One limitation of the proposed strategy is the integrated cost function that combines disparate quantities namely temperatures and power costs. The scaling between these typically competing factors is currently based on trial and error. The magnitude of the cost function itself influences the behavior of the solution in relation to the energy costs and temperature regulation. Furthermore as with all dynamic programming based algorithms, the proposed algorithm also suffers from the curse of dimensionality. However, with model order reduction techniques and efficient LMI based solvers, dynamic programming based solutions can still be feasible for real-time optimization. Alternately, distributed optimization can make real time computation feasible. However, convergence and implementation issues will need to be studied further.

REFERENCES

[1] A. Rantzer, “Relaxed dynamic programming in switching systems,” IEE Proceedings-Control Theory and Applications, vol. 153, no. 5, pp. 567–574, 2006.

[2] W. Zhang, J. Hu, and A. Abate, “On the value functions of the discrete-time switched LQR problem,” Automatic Control, IEEE Transactions on, vol. 54, no. 11, pp. 2669–2674, Nov 2009.

[3] W. Zhang, J. Hu, and A. Abate, “Infinite horizon switched LQR problems in discrete time: A suboptimal algorithm with performance analysis,” IEEE Transactions on Automatic Control, vol. 57, no. 7, pp. 1815–1821, 2012.

[4] D. Gorges, M. Izák, and S. Liu, “Optimal control and scheduling of switched systems,” Automatic Control, IEEETransactionson, vol.56, no.1, pp.135–140, 2011.

[5] Y. Ma, J. Matusko, and F. Borrelli, “Stochastic model predictive control for building hvac systems: Complexity and conservatism,” 2013.

[6] A. Parisio, D. Varagnolo, M. Molinari, G. Pattarello, L. Fabietti, and K. H. Johansson, “Implementation of a scenario-based mpc for hvac systems: an experimental case study,” in The 19th World Congress of the International Federation of Automatic Control, 2014.

[7] D.Kim, E.Cliff, J.E.Braun, J.Borggaard, “Development of control benefit evaluation tool for small commerical buildings”, 2014 ASHRAE/IBPSA-USA Building Simulation Conference

[8] D. Kim, J.E. Braun, J. Borggaard, E. Cliff, and S. Gugercin, “Coupled CFD building envelope model for the Purdue living lab,” in 2nd International High Performance Buildings Conference. 2012.

[9] P. O. Fanger et al., “Thermal comfort. analysis and applications in environmental engineering.” Thermal comfort. Analysis and applications in environmental engineering., 1970.

[10] R. De Dear, “Thermal comfort in practice,” Indoor air, vol. 14, no. s7, pp. 32–39, 2004.

[11] V. Putta, G. Zhu, D. Kim, J. Hu, and J. E. Braun, “Comparative evaluation of model predictive control strategies for a building HVAC system,” in Proceedings of the American Control Conference 2013, 2013, pp. 3461–3467.

[12] V. Putta, G. Zhu, D. Kim, J. Hu, and J. E. Braun, “A Distributed Approach to Efficient Model Predictive Control of Building HVAC Systems,” presented at the International High Performance Buildings Conference, 2012, p. Paper 83.

[13] Y. Ma, G. Anderson, and F. Borrelli, “A distributed predictive control approach to building temperature regulation,” in American Control Conference (ACC), 2011, 2011, pp. 2089–2094.

[14] P.D. Morosan, R. Bourdais, D. Dumur, and J. Buisson, “A dynamic horizon distributed predictive control approach for temperature regulation in multi-zone buildings,” in 2010 18th Mediterranean Conference on Control & Automation (MED), 2010, pp. 622–627.

[15] S. Koehler and F. Borrelli, “Building temperature distributed control via explicit MPC and ‘Trim and Respond’ methods,” in Control Conference (ECC), 2013 European, 2013, pp. 4334–4339.

[16] F. Oldewurtel, A. Parisio, C. N. Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and K. Wirth, “Energy efficient building climate control using stochastic model predictive control and weather predictions,” in American Control Conference (ACC), 2010, 2010, pp. 5100–5105.

Table 1. RTU specifications for restaurant case study			
RTU	**Rated Cooling, kW**	**Rated Power Input, kW**	**COP**
1	52.74	15.06	3.5
2	14.60	5.65	2.58
3	14.60	5.65	2.58
4	14.60	5.65	2.58

Table 2: Performance of the control algorithms over a week			
Algorithm	**Energy consumption, kWh/day**	**Zone averaged RMSE temperature deviation, C**	**Energy savings over baseline, %**
Bang-Bang control	293.6	1.8	Baseline
Centralized MPC (6-hr look-ahead)	283.1	2.1	3.6
Centralized One-shot (6-hr look-ahead)	282.8	2.1	3.7
Distributed MPC (12-hr look-ahead)	292.3	1.9	0.4

Figure I: Layout of case study restaurant

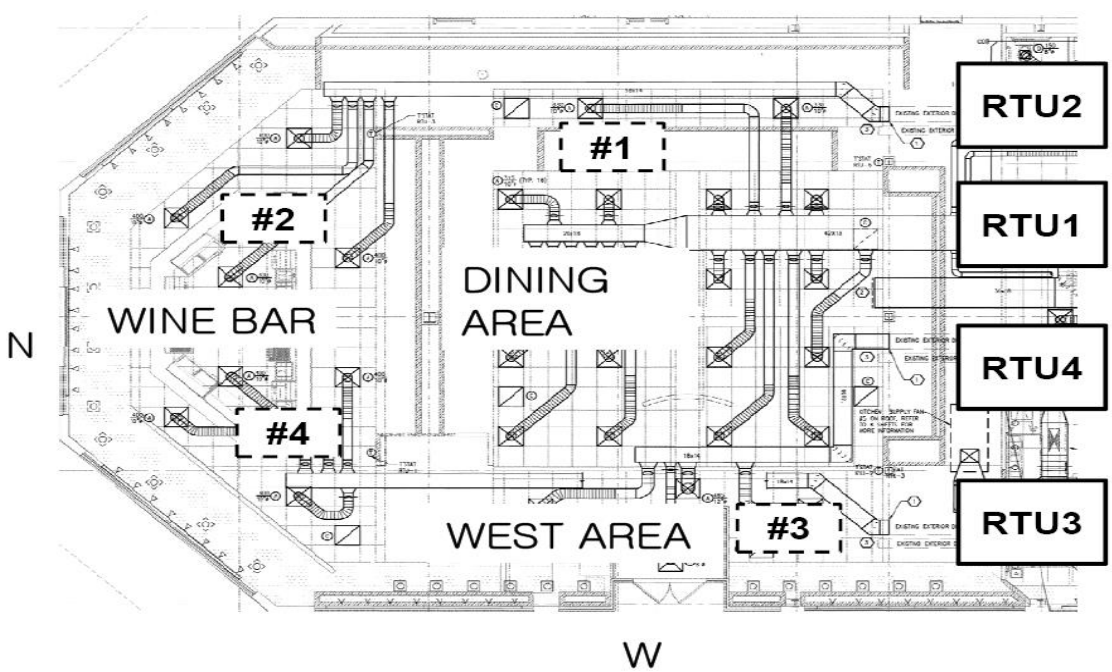

Figure 2: Temperature trajectories of main dining area during validation phase

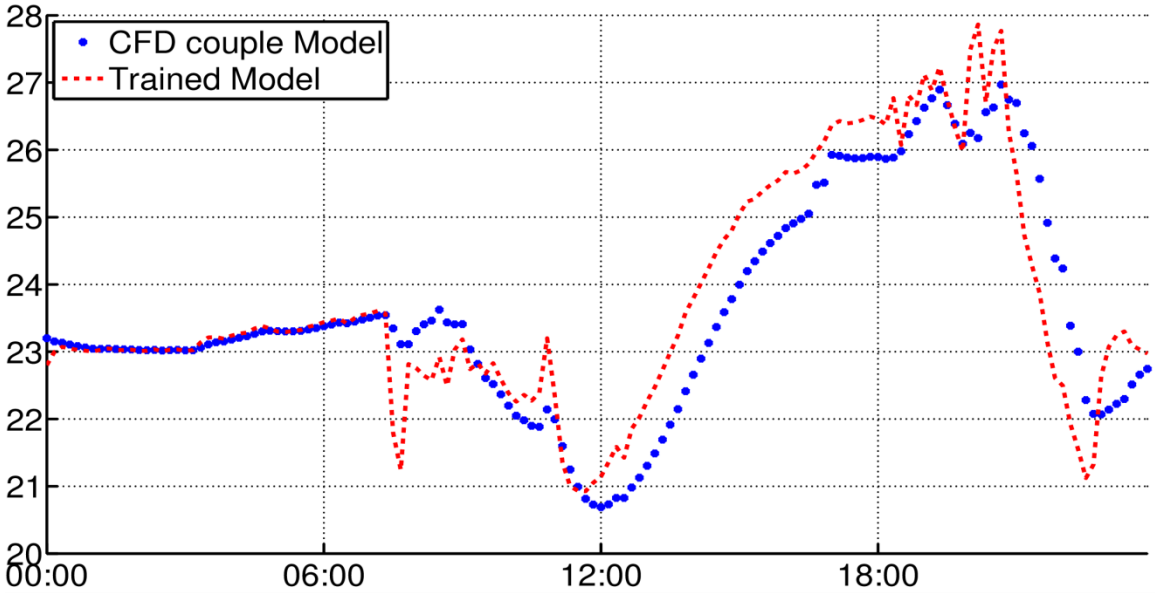

Figure 3a: Typical RTU operation profile under baseline Bang-Bang control

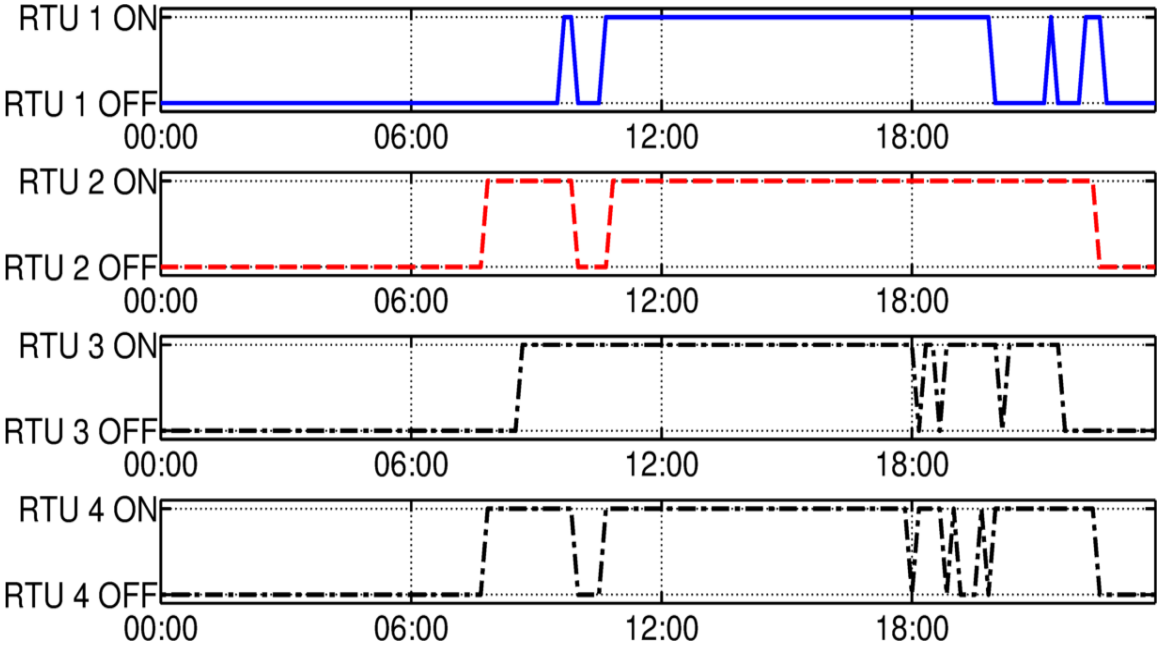

Figure 3b: Typical RTU operation profile under 6hr look-ahead centralized MPC

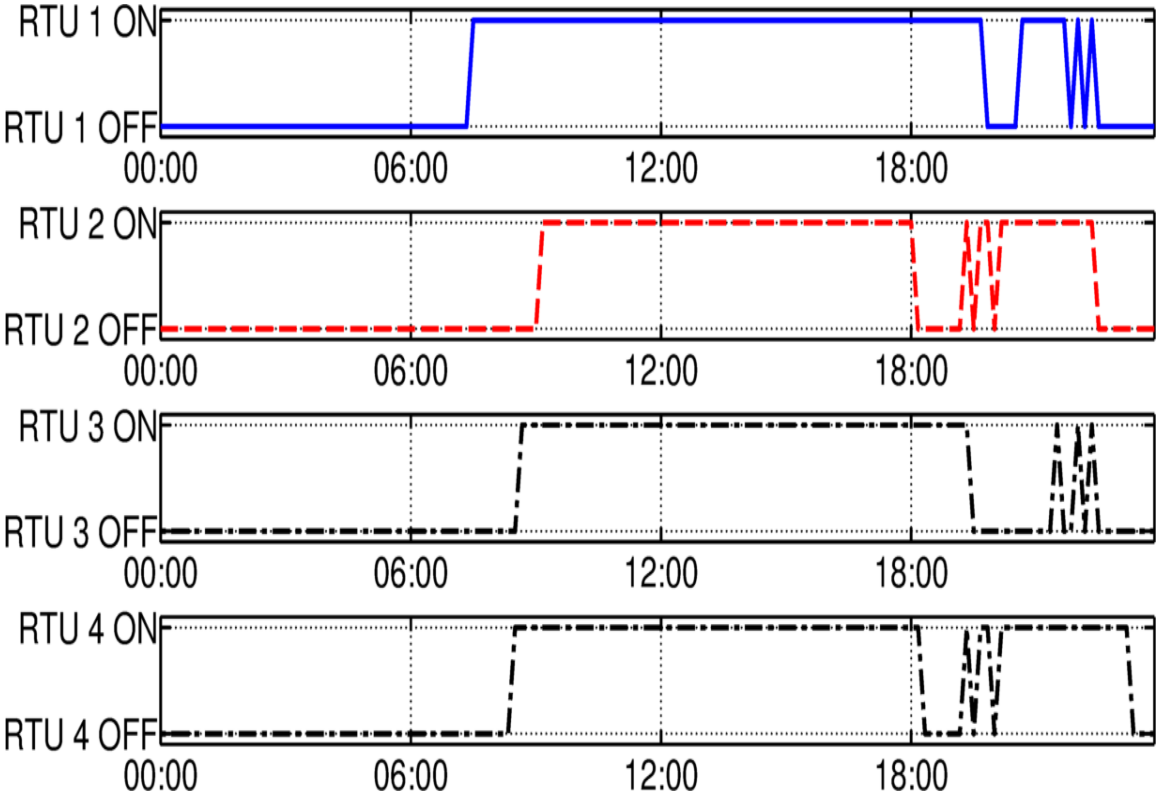

Figure 3c: Typical temperature profile under 6-hr one-shot MPC

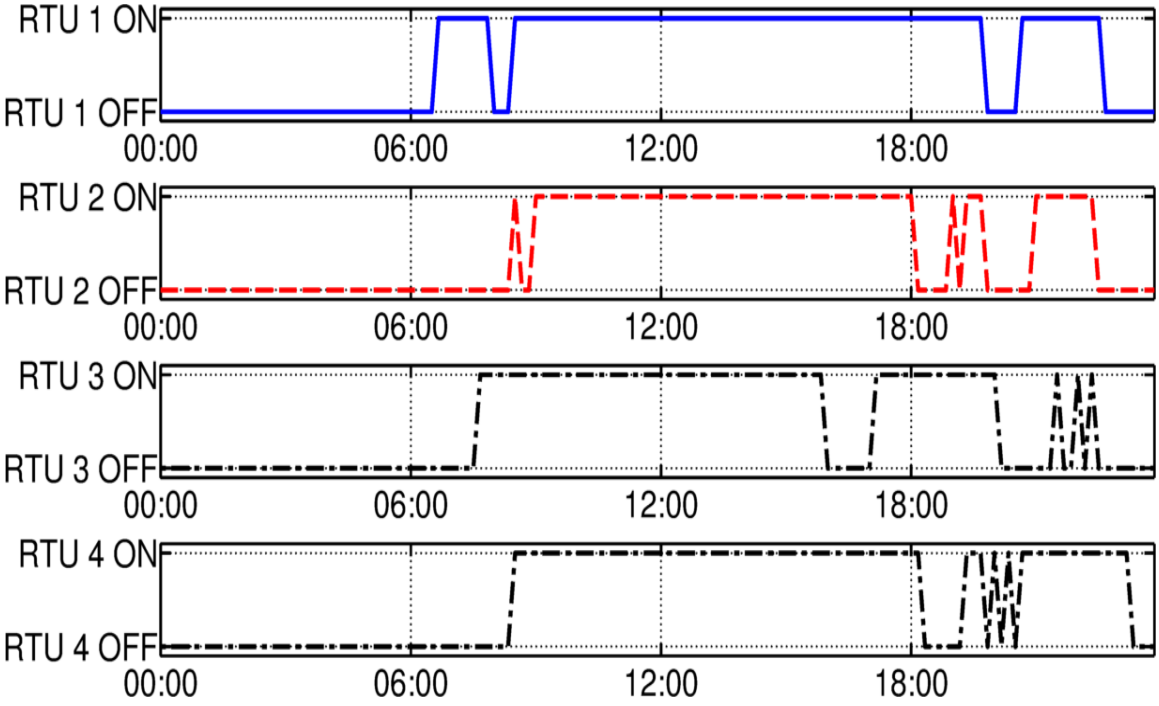

Figure 3d: Typical RTU operation profile under distributed MPC with 12-hour look-ahead

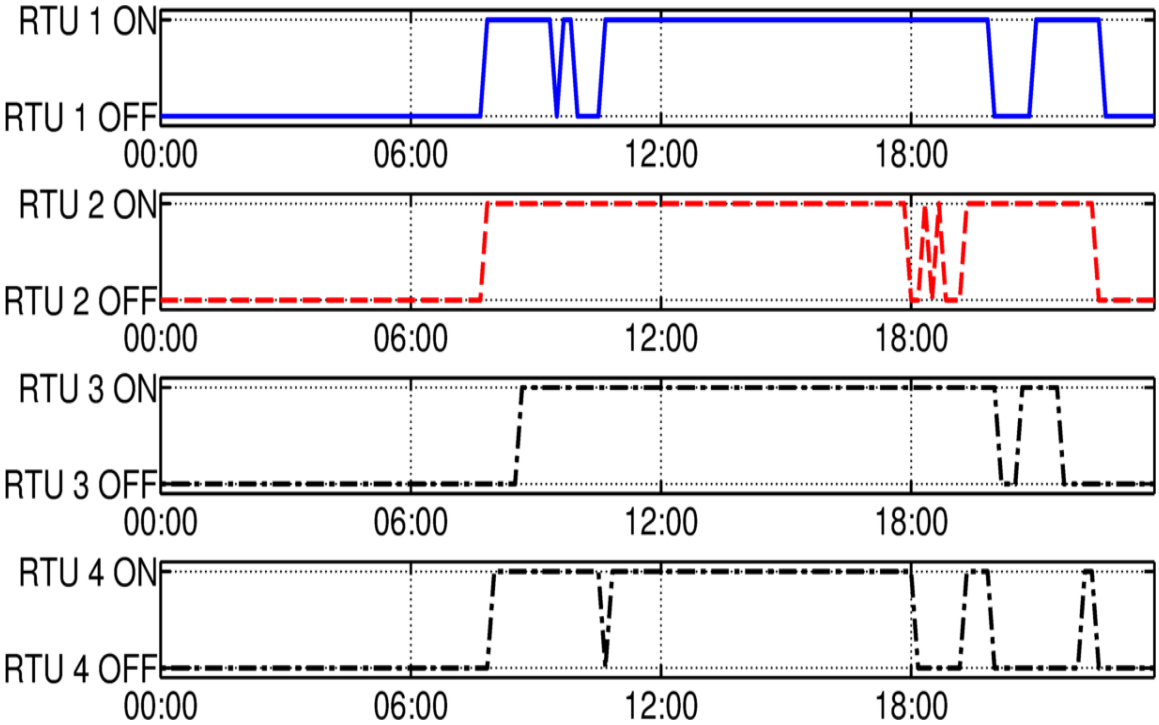

Figure 4: Typical temperature profile of the main dining area under different controls