J.1 Description of IBAL MPC

To control nonlinear system dynamics and discontinuous operation modes (e.g., charging and discharging of active storage) in complex energy systems (e.g., buildings with thermal storage), we propose a two-stage hierarchical predictive controller for optimal management of a complex building energy system, a multizone ice storage plant. This two-stage hierarchical predictive controller is categorized into a high-level controller and a low-level controller. The high-level controller is formulated as a mixed integer linear programming (MILP) problem that schedules cooling energy from different energy sources on a slow-time basis (i.e., 1 hour). The objective is to minimize the energy cost while maintaining sufficient thermal comfort in response to different occupancy profiles, weather conditions and time-of-use energy prices. The low-level controller is formulated as a nonlinear programming (NLP) problem to track the energy and temperature schedules from the top-level controller on a fast-time basis (i.e., 15 minutes). The objective is to minimize the control errors between the schedules and their corresponding measurements, while respecting the HVAC system dynamics that occur on a faster basis.

The two-level hierarchical predictive controller is designed for supervisory control as shown in Figure 1. The control takes the grid signals, disturbance predictions, and system measurements as inputs and generates optimal setpoint schedules for individual devices. The high-level predictive controller adopts an MPC scheme by utilizing the slow dynamics of the ice tank on a slow time (i.e., 1 hour) basis. It is designed as open-loop control for planning the system operation mode, chiller cooling rate, storage heat transfer (e.g., charging and discharging) rate, and zone air temperature setpoints, by minimizing the operation costs for a long time horizon while maintaining thermal comfort in the zones and satisfying the physical constraints of individual equipment. With respect to the optimal schedules from the high-level controller, the low-level controller further generates control policies for different equipment on a fast time (i.e., 15 minutes) basis. For each operation mode other than M0 (system off), a separate Nonlinear Programming (NLP) problem is formulated to coordinate different equipment in that mode to track the optimal plans.

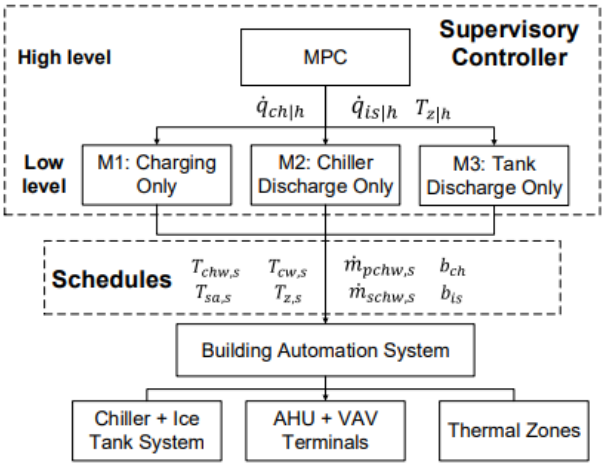

Figure 1 Hierarchical predictive controller for a chiller and ice tank system

The following content describes the details of high-level MPC, low-level MPC, and implementation with the IBAL testbed.

J.1.1 *High-level MPC*

The main purpose of the high-level MPC is to decide which cooling sources should be on and how much cooling energy they should produce to meet the control goals. Eq. (1) - Eq. (4) denote the decision variables for the MPC optimization problem.

$$u_h = [\dot{q}_{ch,c}, \dot{q}_{ch,d}, \dot{q}_{is,d}, \dot{q}_{oa}, \dot{q}_{hvac}, b_{ch,c}, b_{ch,d}, b_{is,d}, z_{P_{ch,d}}, z_{\dot{q}_{ch,d}}, z_{\dot{q}_{is,d}}, \epsilon_T] \quad (1)$$

$$\underline{u}_h = [\underline{\dot{q}}_{ch,c}, \underline{\dot{q}}_{ch,d}, \underline{\dot{q}}_{is,c}, \underline{\dot{q}}_{oa}, \underline{\dot{q}}_{hvac}, 0,0,0, \underline{P}_{ch,d}, \underline{\dot{q}}_{ch,d}, \underline{\dot{q}}_{is,d}, \underline{\epsilon}_T] \quad (2)$$

$$\overline{u}_h = [\overline{\dot{q}}_{ch,c}, \overline{\dot{q}}_{ch,d}, \overline{\dot{q}}_{is,c}, \overline{\dot{q}}_{oa}, \overline{\dot{q}}_{hvac}, 1,1,1, \overline{P}_{ch,d}, \overline{\dot{q}}_{ch,d}, \overline{\dot{q}}_{is,d}, \overline{\epsilon}_T] \quad (3)$$

$$U_h^t = [u_h^t, u_h^{t+1}, \ldots, u_h^{t+N_h-1}] \quad (4)$$

where u_h is the decision variable vector for one step, $\underline{u}_h$ and $\overline{u}_h$ are the lower bound and upper bound, respectively, and U_h^t is the decision variable vector for the MPC problem at time t. $q_{ch,c}$ is charging energy rate by the chiller when charging the ice tank, $\dot{q}_{ch,d}$ is the discharging energy rate by the chiller for cooling the building, $\dot{q}_{is,d}$ is the ice storage discharging rate, $\dot{q}_{oa}$ is the thermal load introduced by maintaining an outdoor air ratio, $\dot{q}_{hvac}$ is cooling energy provided by the HVAC system. The lower bounds for the above four variables are greater than 0, representing the minimum available energy due to physical constraints such as the minimum compressor speed in chillers etc. $b_{ch,c}, b_{ch,d}$ and $b_{is,d}$ are binary variables that represent the chiller in charging mode, the chiller in discharging mode, and ice storage in discharging mode, respectively. $z_{p_{ch,d}}$, $z_{\dot{q}_{ch,d}}$ and $z_{\dot{q}_{is,d}}$ are ancillary variables representing the power consumption of the chiller in cooling the building, the chiller discharging energy rate, and the ice tank discharging rate, respectively, and are used to linearize the otherwise bilinear terms. ϵ_T is a slack variable introduced for zone temperature control, and it has a small positive value.

The high-level formulation is shown below:

$$\min_{U_h^t} \sum_{k=1}^{N_h} [\omega_{h,1} p^{t+k} (P_{cp}^{t+k} + P_{ahu}^{t+k}) \triangle t_h + \omega_{h,2} \epsilon_T^{t+k} + \omega_{h,3} | \triangle u_h^{t+k} |^2] \quad (5)$$

$$\min_{U_h^t} \sum_{k=1}^{N_h} [\omega_{h,1} (P_{cp}^{t+k} + P_{ahu}^{t+k}) \triangle t_h + \omega_{h,2} \epsilon_T^{t+k} + \omega_{h,3} | \triangle u_h^{t+k} |^2] \quad (6)$$

$$\dot{q}_{hvac}^{t+k} = b_{ch,d}^{t+k} \dot{q}_{ch,d}^{t+k} + b_{is,d}^{t+k} \dot{q}_{is,d}^{t+k} + \dot{q}_{oa}^{t+k} \quad (7)$$

$$x^t = x(t) \quad (8)$$

$$T_z^{t+k} = SSM_{z,h}(x^{t+k-1}, u_h^{t+k-1}, d_h^{t+k-1}) \quad (9)$$

$$\underline{T}_z^{t+k} - \epsilon_T^{t+k} \leq T_z^{t+k} \leq \overline{T}_z^{t+k} + \epsilon_T^{t+k} \quad (10)$$

$$\chi^t = \chi(t) \quad (11)$$

$$\chi^{t+k} = \chi^{t+k-1} + \frac{b_{ch,c}^{t+k} \dot{q}_{ch,c}^{t+k} - b_{is,d}^{t+k} \dot{q}_{is,d}^{t+k}}{E_{is,max}} \triangle t_h \quad (12)$$

$$\underline{\chi} \leq \chi^{t+k} \leq \overline{\chi} \quad (13)$$

$$\dot{q}_{oa}^{t+k} = \dot{m}_{oa}^{t+k} C_{pa} (T_{oa}^{t+k} - T_z^{t+k}) \quad (14)$$

$$P_{ch,d}^{t+k} = a_0 + a_1 \dot{q}_{ch,d}^{t+k} \quad (15)$$

$$P_{ch}^{t+k} = b_{ch,c}^{t+k} P_{ch,c}^{t+k} + b_{ch,d}^{t+k} P_{ch,d}^{t+k} \quad (16)$$

$$P_{cp}^{t+k} = P_{ch}^{t+k} + P_{pchw}^{t+k} + (1 - b_{ch,c}^{t+k}) P_{schw}^{t+k} \quad (17)$$

$$P_{ahu}^{t+k} = f_{ahu}(\dot{q}_{hvac}^{t+k}) = \sum_{i=0}^{3} \alpha_i (\dot{q}_{hvac}^{t+k})^i \quad (18)$$

$$b_{ch,c}^{t+k} + b_{ch,d}^{t+k} + b_{is,d}^{t+k} \leq 1 \quad (19)$$

$$\triangle u_h^{t+k} = u_h^{t+k} - u_h^{t+k-1} \quad (20)$$

The objective of the load shifting cases is shown as Eq. (5), and the objective of the energy efficiency and load shedding cases is shown as Eq. (6). The objective has three terms: energy cost or consumption, temperature violation, and skew rates of the control actions. In the energy cost term, p is the energy price, $P_{cp}$is the power consumed by the chiller plant, consisting of chillers, and pumps. P_{ahu} is the power consumed by the AHU fan and $\triangle t_h$ is the control interval of the high-level predictive controller, i.e., 1 hour. In the temperature violation term, the slack variable ϵ_T is used to estimate the temperature violations from given bounds. The skew rate of the control actions $\triangle u_h$ measures how fast the control actions change between adjacent steps, and is defined as Eq. (20). The tradeoffs between these three terms are tuned by weighting factors $\omega_{h,1}$, $\omega_{h,2}$ and $\omega_{h,3}$.

Eq. (7) enforces the energy balance between the cooling system and the thermal zones. It contains two bilinear terms that represent cooling energy provided by discharging the chiller and that provided by discharging the ice tank. This bilinear formulation means the cooling energy provided to the zone is equal to the total amount of cooling energy provided by the chiller and ice tank storage depending on the operation mode. Eq. (8 - 10) describes zone-level equality and inequality constraints, including the zone dynamics model and thermal comfort bounds. Eq. (8) defines the initial states at time t, which are estimated from a linear Kalman Filter by adding Gaussian modeling errors and measurement errors to the discrete time state-space model (SSM). Eq. (9) describes the discrete-time SSN for zone air temperature dynamics. $\underline{T_z}$ and $\overline{T}_z$ are the allowable lower and upper bounds for zone air temperature. Similarly, Eq. (11 - 13) models the storage charging and discharging dynamics in terms of state of charge (SOC) and its physical constraints, where $E_{is,max}$ is the maximum energy that can be stored in the storage.

Eq. (14) models the thermal load introduced by outdoor air. Eq. (15) is a chiller power model for discharging mode as a linear function of the chiller's cooling load, which can be identified from experimental data. Eq. (16) predicts the total power for the chiller by summing up two bilinear terms that represent the power for charging the ice tank and that for providing cooling to the building, respectively. The total power consumed by the whole chiller plant is presented in Eq. (17), which includes the power from the chiller, primary pump and secondary pump. The AHU power is represented as a polynomial equation depending on the total cooling energy provided to the building as shown in Eq. (18). Eq. (19) adds a constraint to represent the physical constraints on the selection of operation modes.

J.1.2 *Low-level MPC*

The low-level predictive controller aims to optimally operate the system under given operation modes while respecting the schedules from the high-level MPC. The low-level predictive controller operates at a faster time step (e.g., 15 min) to respect the system states (i.e., zone temperature and tank SOC) and schedules (i.e., energy transfer rate from different cooling sources) optimized from the high-level controller. Except for M0 where the system is off, and M1 where the system is under charging mode with a predefined fixed schedule, we established an optimal control problem for each operation mode.

M2: Chiller Discharging Mode

When $b_{ch,d}$ = 1 and $b_{is,d}$ = 0, mode M2 is activated so that only the chiller is on to provide cooling to the building. To respect the high-level controller, the following problem is formulated:

$$\min_{U_l^t}[\omega_1|\epsilon_T^{t+N_t}|^2 + \sum_{k=1}^{N_t}\omega_2|\epsilon_q^{t+k}|^2] \quad (21)$$

$$u_l = [T_{cws}, T_{chws}, T_{ma}, \triangle T_{cc}, \dot{m}_{schw}, \dot{m}_{sa}, \dot{q}_{ch}, \epsilon_T, \epsilon_q] \quad (22)$$

$$U_l^t = [u_l^t, u_l^{t+1}, \ldots, u_l^{t+N_l-1}] \quad (23)$$

$$x^t = x(t) \quad (24)$$

$$T_z^{t+k} = SSM_{z,l}(x^{t+k-1}, u_l^{t+k-1}, d_l^{t+k-1}) \quad (25)$$

$$P_{ch}^{t+k} = \alpha_0 + \alpha_1 T_{chw}^{t+k} + \alpha_2 (T_{chw}^{t+k})^2 + \alpha_3 T_{cw}^{t+k} + \alpha_4 (T_{cw}^{t+k})^2 + \alpha_5 T_{chw}^{t+k} T_{cw}^{t+k} + \alpha_6 (\dot{q}_{ch}^{t+k}) + \alpha_7 (\dot{q}_{ch}^{t+k})^2 \quad (26)$$

$$P_{pchw}^{t+k} = \overline{P}_{pchw} \quad (27)$$

$$P_{schw}^{t+k} = \sum_{i=0}^{3}\beta_i (\dot{m}_{schw}^{t+k})^i \quad (28)$$

$$P_{cp}^{t+k} = P_{ch}^{t+k} + P_{pchw}^{t+k} + P_{schw}^{t+k} \quad (29)$$

$$P_{ahu}^{t+k} = \sum_{i=0}^{3}\gamma_i (\dot{m}_{sa}^{t+k})^i \quad (30)$$

$$\dot{q}_{ch}^{t+k} = \dot{m}_{sa}^{t+k} C_{pa} \triangle T_{cc}^{t+k} \quad (31)$$

$$\dot{q}_{ch}^{t+k} = \dot{m}_{schw}^{t+k} C_{pw} (T_{chwr,coi}^{t+k} - T_{chws}^{t+k}) \quad (32)$$

$$\dot{m}_{oa}^{t+k} T_{oa}^{t+k} + (\dot{m}_{sa}^{t+k} - \dot{m}_{oa}^{t+k}) T_z^{t+k} - \dot{m}_{sa}^{t+k} T_{ma}^{t+k} = 0 \quad (33)$$

$$T_{sa}^{t+k} = T_{ma}^{t+k} - \triangle T_{cc}^{t+k} \quad (34)$$

$$T_{chwr,coi}^{t+k} \leq T_{sa}^{t+k} - \zeta_1 \quad (35)$$

$$T_{chws}^{t+k} \leq T_{sa}^{t+k} - \zeta_2 \quad (36)$$

$$\frac{\dot{q}_{ch,cap}^{t+k}}{\bar{\dot{q}}_{ch}} = \delta_0 + \delta_1 T_{chws}^{t+k} + \delta_2 (T_{chws}^{t+k})^2 + \delta_3 T_{cws}^{t+k} + \delta_4 (T_{cws}^{t+k})^2 + \delta_5 T_{chws}^{t+k} T_{cws}^{t+k} \quad (37)$$

$$PLR_{min}\dot{q}_{ch,cap}^{t+k} \leq \dot{q}_{ch}^{t+k} \leq PLR_{max}\dot{q}_{ch,cap}^{t+k} \quad (38)$$

$$-\epsilon_T \leq T_{z\leftarrow h}^{t} - T_z^{t+N_t} \leq \epsilon_T \quad (39)$$

$$-\epsilon_q^{t+k} \leq \dot{q}_{ch,d\leftarrow h}^{t} - \dot{q}_{ch}^{t+N_t} \leq \epsilon_q^{t+k} \quad (40)$$

$$\sum_{k=1}^{M} (P_{cp}^{t+k} + P_{ahu}^{t+k}) \triangle t_l \leq (P_{cp\leftarrow h}^{t} + P_{ahu\leftarrow h}^{t}) \triangle t_h \quad (41)$$

The control objective as shown in Eq. (21) is to respect the optimal schedules generated from the high-level controller, and thus to minimize the zone temperature differences and the chiller cooling rate differences between the high-level schedules and the low-level fulfillment. The minimal differences are achieved by searching the decision variables u_l over the future time span $[t, t + N_l]$. For each time step, the decision vector ul defined as Eq. (22) includes the condenser water supply temperature T_{cws}, chilled water supply temperature T_{chws}, air-side temperature difference across the cooling coil $\triangle T_{cc}$, secondary pump mass flow rate $\dot{m}_{schw}$, supply air mass flow rate $\dot{m}_{sa}$, chiller cooling rate $\dot{q}_{ch}$, slack tolerance ϵ_T for zone temperature control, and ϵ_q for chiller cooling rate balance.

Eq. (24) and Eq. (25) describe the low-level discrete-time SSM for zone air temperature dynamics. Eqs. (26-30) model the power of the major cooling system device, such as the chiller, primary chilled water pump and secondary chilled water pump, and AHU, where $\overline{P}_{pchw}$ is the nominal power by the CHWP, and the Greek symbols α, β, γ are the identified model parameters from experimental data. Eqs. (31-36) add constraints for the steady-state energy balance between the air-side and water-side of the cooling coil, where ζ_1 is the allowable temperature difference between the chilled water return temperature and the supply air temperature, and ζ_2 is the approach temperature of the cooling coil.

Eqs. (37-38) bound the chiller cooling rate based on operational conditions such as CHWST and allowable part load ratio (PLR), where $\dot{q}_{ch,cap}$ is the chiller's available cooling capacity, $\overline{\dot{q}}_{ch}$ is the nominal capacity, and δ represents the curve coefficients calibrated from experimental data. Eqs. (39 - 41) are inequality constraints that force the low-level controller to respect the high-level schedules in terms of zone air temperature, chiller cooling rate and total power consumption, where $T_{z\leftarrow h}^{t}$ $\dot{q}_{ch,d\leftarrow h}^{t}$, $P_{cp\leftarrow h}^{t}$ and $P_{ahu\leftarrow h}^{t}$ are the zone temperature schedule, chiller cooling rate schedule, chiller plant power schedule and AHU power schedule from the high-level controller generated at time t, respectively.

M3: Ice Tank Discharging Mode

This mode is activated when $b_{is,d}$ = 1 and $b_{ch,d}$ = 0. To respect the high-level controller, the following problem is established with the objective of minimizing the termination cost of zone air temperature and ice tank SOC, and tracking errors for ice tank discharging rate schedules. The decision variables are listed as Eqs. (43) and (23), where $\dot{q}_{is}$ is the low-level ice tank discharging rate, and ϵ_χ is the slack variable for the terminal tank SOC.

$$\min_{U_l^t} [\omega_1 |\epsilon_T^{t+N_t}|^2 + \omega_2 |\epsilon_\chi^{t+N_t}|^2 + \sum_{k=1}^{N_t} \omega_3 |\epsilon_q^{t+k}|^2] \quad (42)$$

$$u_l = [T_{chws}, \triangle T_{cc}, \dot{m}_{schw}, \dot{m}_{sa}, \dot{q}_{is}, \epsilon_T, \epsilon_\chi, \epsilon_q] \quad (43)$$

$$\chi^t = \chi(t) \quad (44)$$

$$\chi^t = \chi^{t+k-1} - \frac{\dot{q}_{is,d}^{t+k} \triangle t_l}{E_{is,max}} \quad (45)$$

$$P_{cp}^{t+k} = P_{pchw}^{t+k} + P_{schw}^{t+k} \quad (46)$$

$$\dot{q}_{is}^{t+k} = \dot{m}_{sa}^{t+k} C_{pa} \triangle T_{cc}^{t+k} \quad (47)$$

$$\dot{q}_{is}^{t+k} = \dot{m}_{schw}^{t+k} C_{pw} (T_{chwr,coi}^{t+k} - T_{chws}^{t+k}) \quad (48)$$

$$\dot{q}_{is,cap}^{t+k} = \sum_{i=0}^{3} \eta_i (\dot{\chi}^{t+k})^i \quad (49)$$

$$PLR_{min} \dot{q}_{is,cap}^{t+k} \leq \dot{q}_{is}^{t+k} \leq PLR_{max} \dot{q}_{is,cap}^{t+k} \quad (50)$$

$$-\epsilon_{\chi}^{t+k} \leq \chi_{\leftarrow h}^{t} - \chi^{t+N_l} \leq \epsilon_{\chi}^{t+N_l} \quad (51)$$

$$-\epsilon_{q}^{t+k} \leq \dot{q}_{is,d\leftarrow h}^{t} - \dot{q}_{is}^{t+\mathrm{k}} \leq \epsilon_{q}^{t+k} \quad (52)$$

Eqs. (44-45) are the ice tank SOC dynamics, and Eq. (46) is the total power for the hydraulic system. Eq. (47-48) are steady-state heat balance equations for the cooling coil under M3. The ice tank time-variant discharging capacity is modeled as a polynomial equation of its current SOC as described in Eq. (49). This controller should also respect the high-level optimal terminal conditions for the zone air temperature (Eq.(39)) and ice tank SOC (Eq.(51)), and the optimal schedules for ice tank discharging rate (Eq. (52)) and system power (Eq. (41)).

J.1.3 *Zone Temperature State Space Model*

A dynamic zone air temperature predictor for the aggregated zone is needed when the zone air temperature is used for the high-level and low-level MPC. A lumped RC thermal network model is used to predict the average zone temperature for each of the two zones that are served by the same AHU. The building thermal load model is sketched in Figure 2. The resulting state space model (SSM) based on energy conservation is shown in Eqs. (53-55).

$$C_z \frac{dT_z}{dt} = \frac{T_{oa} - T_z}{R_g} + \frac{T_{w,i} - T_z}{R_i} + \dot{q}_{conv,i} + \dot{q}_{hvac} \quad (53)$$

$$C_{w,e} \frac{dT_{w,e}}{dt} = \frac{T_{oa} - T_{w,e}}{R_e} + \frac{T_{w,i} - T_{w,e}}{R_w} + \dot{q}_{sol,e} \quad (54)$$

$$C_{w,i} \frac{dT_{w,i}}{dt} = \frac{T_{w,e} - T_{w,i}}{R_w} + \frac{T_z - T_{w,i}}{R_i} + \dot{q}_{rad,i} \quad (55)$$

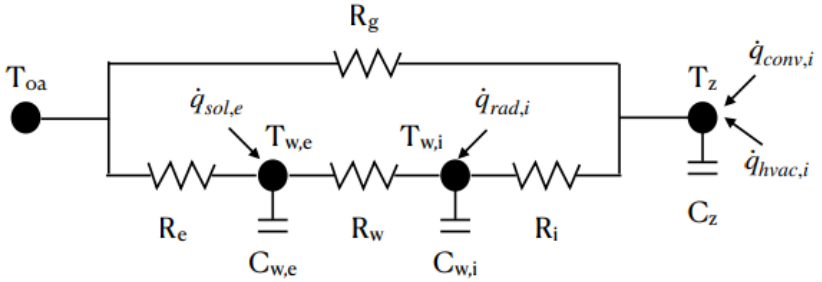

Figure 2 RC-based thermal zone model. R_g represents the thermal resistance of windows. The walls are separated into two layers: $C_{w,e}$ and $C_{w,i}$ are thermal capacitance of the exterior layer and interior layer respectively. The thermal resistance between $C_{w,e}$ and $C_{w,i}$ is modeled by R_w, while R_e and R_i capture the thermal resistance associated with heat convection across the

exterior and interior wall surfaces respectively. C_z is the thermal capacitance in the zone. The model inputs are outside air temperature T_{oa}, external solar load $\dot{q}_{sol,e}$, internal radiative heat gain $\dot{q}_{rad,i}$, internal convective heat gain $\dot{q}_{conv,i}$, and HVAC heat rate $\dot{q}_{hvac,i}$. The model output is the zone air temperature T_z.

J.1.4 *Implementation with IBAL*

The two-stage hierarchical predictive controller proposed above is a generalized control method for VAV AHU systems. There is a gap between this control strategy and the actual IBAL system, specifically that the controller is considering the case of a single zone, whereas the IBAL system has four zones controlled by two AHUs. In this regard, when applying the MPC in practice, the weighted average zone air temperature is used as the measurement value input to the MPC. Then the MPC performs subsequent operations based on this weighted average zone air temperature. The weighting ratio is determined based on the maximum cooling load of the four zones during the peak period. Specifically, the four zones in IBAL are Conference Room (serviced by AHU1, minimum flow rate = 460 CFM), Enclosed Office 1 (serviced by AHU1, minimum flow rate = 200 CFM), Open Office (serviced by AHU2, minimum flow rate = 200 CFM), and Enclosed Office 2 (serviced by AHU2, minimum flow rate = 200 CFM). Based on simulation, the weighting ratio of the four zones in the peak period is [0.31, 0.18, 0.33, 0.18]. In addition to this, since the MPC only considers one zone, the supply air mass flow rate setpoint output from the MPC is an aggregated mass flow rate setpoint. This aggregated setpoint is assigned to the AHUs based on the same weighting ratio. The lower bound of the $\dot{m}_{sa}$ is set to 1500 CFM to account for the fact that after assigning, the air supply also needs to meet the minimum air supply requirement.

J.2 The Limitation of IBAL MPC

The MPC strategies utilized in this study represent one of many potential approaches. Specifically, the formulation developed by the TAMU team was designed with an intentional focus on generality and applicability in diverse settings. It is geared towards optimizing setpoints at an aggregated level and assuming limited knowledge about the dynamic of each individual zone and each AHU. While this approach may not be the most optimal, it serves as a practical example of how MPC can be feasibly applied in real-time.

The results illustrated in **Error! Reference source not found.** through **Error! Reference source not found.** indicate that MPC cases typically demonstrate reduced energy consumption, better flexibility, and lower electricity costs in comparison to the default cases with RBC. However, these benefits are accompanied by certain trade-offs. Figure 3 shows the zone air temperature of each individual zone from Tucson_Shift_Default and Tucson_Shift_MPC, respectively. In the Tucson_Shift_Default case, where RBC was implemented, zone temperatures were consistently kept below the cooling setpoint. However, this occasionally resulted in over-cooling due to the minimum flow rate setpoint. On the other hand, in the MPC case, the zone temperatures of Conference Room Mid 2 and Open Office Mid 1 were not maintained at the setpoint. This discrepancy stems from the use of an aggregated zone Resistance-Capacitance (RC) model in the MPC. During model calibration, the aggregated zone temperature was calculated as a weighted average across four zones, with the weights assigned based on each zone's proportion of the maximum cooling load. These weights, however, do not capture the actual weight of each zone during a daily simulation due to the variations in weather and internal load schedules. For example, Enclosed Office Mid 2 has higher heat gain in the morning than in the afternoon because of the east-facing wall. Consequently, these dynamically changing weights lead to less accurate

predictions of zone temperature setpoints, subsequently resulting in less accurate predictions of system cooling load. In observed cases, the cooling load was often underpredicted, making it challenging to maintain the cooling setpoints in certain zones. Despite these limitations, the practice of setting weights based on the proportion of the maximum cooling load stems from an assumption of limited operational knowledge. The maximum cooling load, or the design load, is often the most readily accessible information for a building operator, hence it is used as a basis for weight determination.

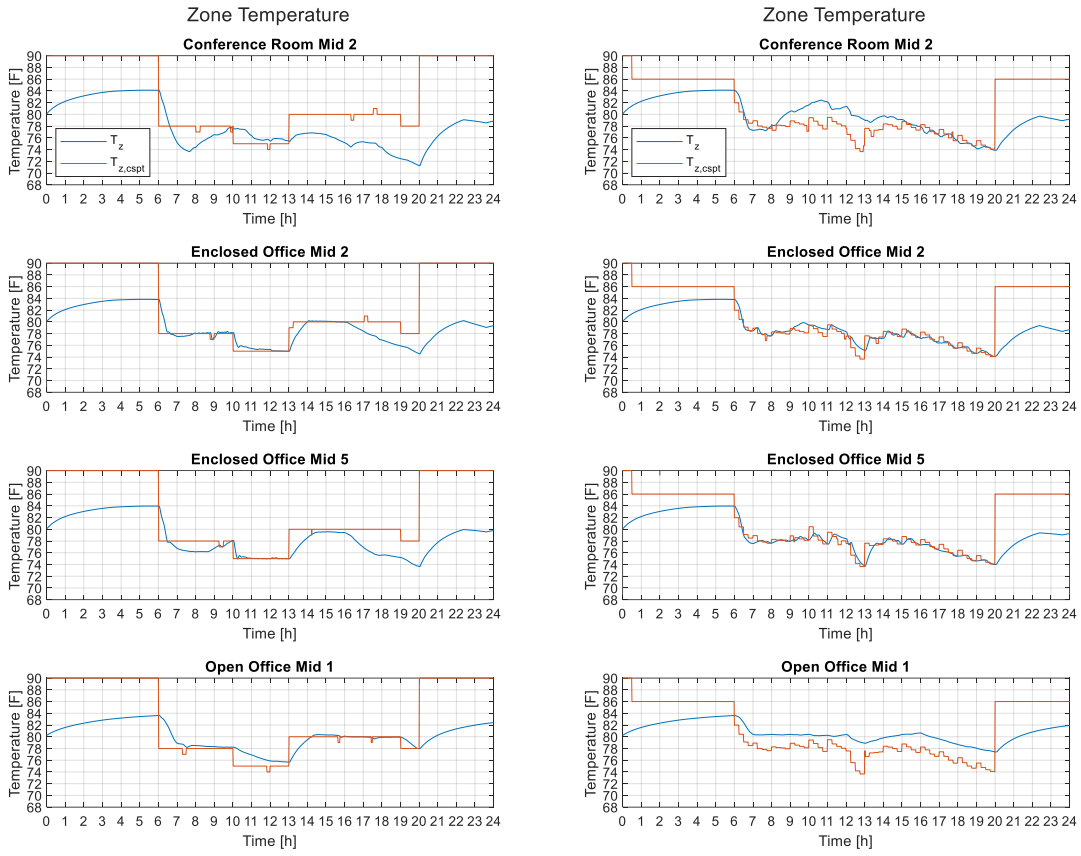

Figure 3 Air temperature of each zone from Tucson_Shift_Default (left) and Tucson_Shift_MPC (right)

In analyzing the MPC results, it was also observed that the minimum flow rate setpoints for the VAV units were not always met. Figure 4 (left) shows the VAV flow rate in each zone for a RBC case, i.e., Tucson_Shift_Default. The blue line in the figure shows the measured flow rates of each VAV box. The orange line indicates the actual ventilation requirements calculated based on the 20 CFM/person rate and the occupancy at each time step. The yellow line represents the minimum flow rate setpoint. Note that the actual maximum ventilation requirements for the two enclosed offices were 40 CFM and they are not the same as the minimum flow rate setpoint. This is because the hardware emulators require a minimum of 200 CFM air flow to accurately emulate zone conditions. The plots confirm that under RBC, the minimum flow rate setpoints were consistently achieved. Contrastingly, Figure 4 (right) shows the VAV flow rate for each zone in an MPC case,

i.e., Tucson_Shift_MPC. Here, adherence to the designed minimum flow rates under the MPC was less stringent. This was mainly evident in Conference Room Mid 2.

In the MPC, the aggregated flow rate setpoint was distributed using the same weights as those used for the weighted temperature model. The minimum aggregated flow rate, 1500 CFM, is sufficient to satisfy the minimum flow rate of each zone if the weights are 0.31, 0.18, 0.33, 0.18 for Conference Room Mid2, EnclosedOfficeMid2, EnclosedOfficeMid5, and OpenOfficeMid1, respectively. However, due to the changing nature of these weights throughout the day, the flow rate distributed to an AHU might occasionally fall short of requirements. Nevertheless, despite the measured VAV flow rates not aligning with the minimum flow rate setpoints, they were generally found to be adequate for the occupancy levels during each time step of the test cases in this project.

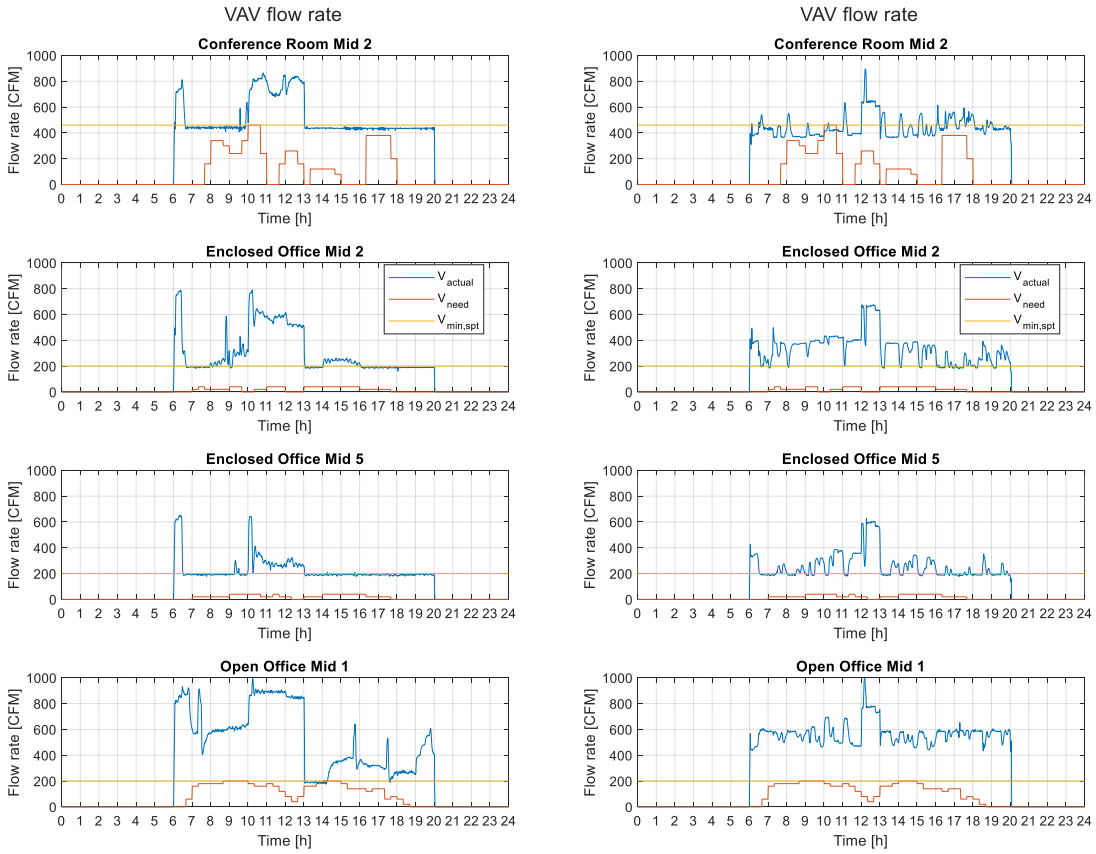

Figure 4 VAV flow rate in each zone from Tucson_Shift_Default (left) and Tucson_Shift_MPC (right)

In conclusion, the analysis of the MPC results within this study reveals an overall positive performance, highlighting the feasibility and effectiveness of the MPC approach. Despite certain limitations, such as unmet cooling setpoint and occasional deviations from minimum flow rate setpoints, the MPC strategies implemented have successfully demonstrated key benefits. These include reduced energy consumption, enhanced flexibility, and lower electricity costs compared to traditional RBC strategies.