

Bilinear Model Predictive Control of a HVAC System Using Sequential Quadratic Programming

Anthony Kelman* Francesco Borrelli**

* Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1740 USA (e-mail: kelman@berkeley.edu)

** Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1740 USA (e-mail: fborrelli@berkeley.edu)

Abstract: We study the problem of heating, ventilation, and air conditioning (HVAC) control in a typical commercial building. We propose a model predictive control (MPC) approach which minimizes energy use while satisfying occupant comfort constraints. A sequential quadratic programming algorithm is used to efficiently solve the resulting bilinear optimization problem. This paper presents the control design approach and the procedure for computing its solution. Extensive numerical simulations show the effectiveness of the proposed approach. In particular, the MPC is able to systematically reproduce a variety of well-known commercial solutions for energy savings, which include demand response, “economizer mode” and precooling/preheating.

1. INTRODUCTION

The building sector consumes about 40% of the energy used in the United States and is responsible for nearly 40% of greenhouse gas emissions, see McQuade (2009). It is therefore economically, socially and environmentally significant to reduce the energy consumption of buildings. Previous work by Henze et al. (2004); Ma et al. (2010); Oldewurtel et al. (2010) has evaluated the energy saving potential of model predictive control (MPC) for heating ventilation and air conditioning (HVAC) in buildings.

The main idea of MPC is to use a model of the plant to predict the future evolution of the system, see Borrelli (2003); Mayne et al. (2000). At each discrete sampling time, an open-loop optimal control problem is solved over a finite horizon starting from the current state. The optimal command signal is applied to the process only during the first sampling interval. At the next time step a new optimal control problem based on new measurements of the state is solved over a shifted horizon. Model predictive control has become the accepted standard in the process industry for solving complicated constrained multivariable control problems, see Qin and Badgwell (2003). The success of MPC is largely due to its ability to simply and effectively handle hard constraints on states and control inputs.

Section 2 of this paper introduces the common type of HVAC system we focus on, then derives low-order models for the temperature dynamics and energy costs. Section 3 outlines the MPC control algorithm and presents a tailored sequential quadratic programming (SQP) method which exploits the bilinear structure of the optimal control problem. Simulation results presented in Section 4 show good performance and computational tractability of the resulting scheme. Additionally, the model predictive controller exhibits desirable control behaviors that resemble modern advanced heuristic strategies for HVAC control, reproducing those strategies in a systematic manner.

2. SYSTEM MODEL

The system considered in this work, known as variable air volume (VAV) with reheat, is depicted in Fig. 1. This type of HVAC system provides cooling using a variable speed supply fan and a cooling coil (water-to-air heat exchanger) located in a central air handling unit (AHU). A cold water valve controls the cooling coil outlet air to a setpoint temperature, and the supply fan distributes the cold air to every thermal zone served by the same AHU. The AHU supply fan speed determines the total flow rate to all zones. The supply air flows from the AHU into each thermal zone through a “VAV box” which consists of a damper and a heating coil. The damper regulates the flow rate to the thermal zone and the heating coil reheats the supply air to a higher temperature, if required.

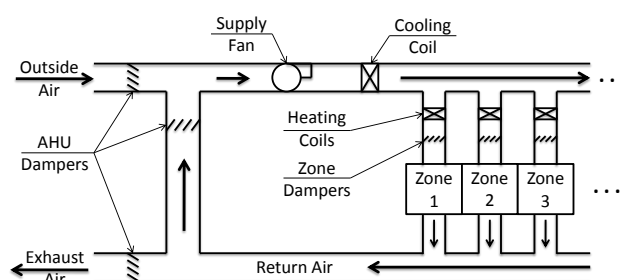


Fig. 1. VAV with reheat HVAC system schematic

The mixed exit air from all zones recirculates back to the AHU through a return duct. A set of dampers in the AHU controls the mix of outside ambient and recirculated air used as the source for supply air. Usually the return air will be cooler than ambient temperature on a hot day when cooling is required, or warmer than ambient on a cold day when heating is required. So conventional practice is to

recirculate as much air as possible, while maintaining a minimum fraction of fresh air for acceptable indoor air quality. However, the opposite scenario of cooling when return air is warmer than ambient (or heating when return air is cooler than ambient) can occur, and in that case using 100% outside air consumes the least total coil energy. This is known as “economizer” operation.

In order to develop a low-order control-oriented building thermal model, we make the following assumptions.

- A1 The average air temperature dynamics of the thermal zones can be reasonably approximated as first-order. We therefore combine the thermal capacitance of the air, walls, furnishings, etc. of zone i into a single lumped parameter denoted $(mc)_i$.
- A2 Humidity is not explicitly modeled.¹
- A3 All dynamics except those of the thermal zones are neglected. Actuators are assumed to instantly meet their control setpoints.
- A4 Predictions of outside ambient temperature T_{oa} and the thermal loads \dot{Q}_i in each zone due to occupants, equipment, and all heat transfer to or from ambient and other zones are known in advance.²

2.1 Zone Temperature Dynamics

A first order energy balance gives the following continuous time system dynamics for the temperature T_{zi} of zone i .

$$(mc)_i \frac{d}{dt} T_{zi} = \dot{Q}_i + \dot{m}_{zi} c_p (T_{si} - T_{zi}). \quad (1)$$

where c_p is the specific heat capacity of air, and the control inputs are the mass flow rates \dot{m}_{zi} and supply temperatures T_{si} to each zone.

If we neglect heat transfer by radiation, the thermal loads \dot{Q}_i are affine functions of the zone temperatures,

$$\dot{Q} = \begin{bmatrix} \dot{Q}_1 \\ \vdots \\ \dot{Q}_n \end{bmatrix} = R \begin{bmatrix} T_{z1} \\ \vdots \\ T_{zn} \end{bmatrix} + \dot{Q}_{\text{offset}},$$

where n is the number of thermal zones served by the same AHU, R is a symmetric $n \times n$ matrix and \dot{Q}_{offset} is an $n \times 1$ time-varying vector. The elements of R are given by

$$R_{ij} = \begin{cases} UA_{ij} & i \neq j \\ -UA_{oi} - \sum_{k=1}^{i-1} UA_{ik} - \sum_{k=i+1}^n UA_{ik} & i = j \end{cases}$$

where $UA_{ij} = UA_{ji}$ is the heat transfer coefficient times area between zones i and j . Zone i and zone j have no heat transfer between them when $UA_{ij} = UA_{ji} = 0$. The heat transfer coefficient times area between zone i and specified external temperatures (such as T_{oa}) is denoted UA_{oi} .

¹ Humidity and latent heat can be important factors affecting HVAC design and actual energy use, but they are not typically directly measured or controlled. We will deal with this thermodynamic simplification indirectly: throughout this entire work, every temperature is meant to represent an equivalent dry air temperature, i.e. the temperature at which dry air would have the same specific enthalpy as the actual moist air mixture.

² This is a strong assumption, since load and weather predictions are not perfectly accurate. We are investigating extensions of this MPC scheme to robustly account for prediction and model uncertainty.

We can use a weighted undirected graph representation in order to describe the coupling between zones. In that case, $-R$ is equal to the weighted Laplacian matrix of the zone connectivity graph plus a diagonal matrix.

The compact form of (1) for all zones together is then

$$M \frac{d}{dt} T_z = R T_z + \dot{Q}_{\text{offset}} + c_p \dot{m}_z \circ (T_s - T_z), \quad (2)$$

where $M = \text{diag}((mc)_1, \dots, (mc)_n)$, $T_z = [T_{z1}, \dots, T_{zn}]'$, $\dot{m}_z = [\dot{m}_{z1}, \dots, \dot{m}_{zn}]'$, $T_s = [T_{s1}, \dots, T_{sn}]'$, and \circ denotes the elementwise product.

Assuming \dot{m}_z and T_s are zero-order held at sample rate Δt , we discretize (2) using the trapezoidal method to obtain

$$M \frac{T_z^+ - T_z}{\Delta t} = R \frac{T_z^+ + T_z}{2} + \frac{\dot{Q}_{\text{offset}}^+ + \dot{Q}_{\text{offset}}}{2} + c_p \dot{m}_z \circ \left(T_s - \frac{T_z^+ + T_z}{2} \right), \quad (3)$$

where T_z^+ and $\dot{Q}_{\text{offset}}^+$ denote the corresponding values at the next discrete time step $t + \Delta t$. We choose the trapezoidal discretization as a compromise between simplicity and stability (which some methods lack when Δt is large).

2.2 Air Handling Unit Model

Although we have neglected the dynamics of the AHU, we must introduce additional control inputs and constraints to capture recirculation behavior and the energy cost.

Table 1. Additional variables for AHU model

Symbol	Description	Units
d_r	Fraction of supply flow recirculated from zones	dimensionless
T_r	Temperature of mixed return air from zones	$^{\circ}\text{C}$
T_m	AHU cooling coil inlet air temperature	$^{\circ}\text{C}$
T_c	AHU cooling coil outlet air temperature setpoint	$^{\circ}\text{C}$

Assuming there is no heat loss or gain in the return duct, T_r is a flow-weighted average of the zone temperatures,

$$T_r = \frac{\sum_{i=1}^n (\dot{m}_{zi} T_{zi})}{\sum_{i=1}^n \dot{m}_{zi}}. \quad (4)$$

An energy balance of the AHU then gives

$$T_m = (1 - d_r) T_{oa} + d_r T_r. \quad (5)$$

2.3 Cost Function

We will consider energy use of the cooling and heating coils in terms of air-side thermal power, which we can directly calculate from the models in the previous section. We represent all of the operating characteristics of the cold and hot water circuits with two constant parameters: efficiency η_h for the hot side, and coefficient of performance η_c for the cold side. So power used by the heating coils and cooling coil are given respectively by P_h and P_c where

$$P_h = \frac{c_p}{\eta_h} \sum_{i=1}^n (\dot{m}_{zi} (T_{si} - T_c)), \quad (6)$$

$$\text{and } P_c = \frac{c_p}{\eta_c} \left(\sum_{i=1}^n \dot{m}_{zi} \right) (T_m - T_c). \quad (7)$$

Remark 1. A higher-fidelity model would include ancillary equipment such as water pumps and cooling towers, and a detailed water-side energy balance.

The electrical power P_f used by the supply fan is equal to

$$P_f = \left(\sum_{i=1}^n \dot{m}_{zi} \right) \frac{\Delta p}{\rho \eta_f},$$

where Δp is the pressure difference across the fan, ρ is the air density, and η_f is the efficiency of the fan. If all zone dampers were held at fixed positions, incompressible flow would give $\Delta p \propto (\sum \dot{m}_{zi})^2$. However, when zone dampers open, the pressure drop for a given flow rate decreases. As an average trend direction, increasing total flow will correspond to more-open zone dampers. So we expect pressure drop to increase more slowly than total flow squared. For simplicity we restrict our model to polynomial form, so we take $\Delta p = \rho \eta_f \kappa_f \sum \dot{m}_{zi}$. Recorded data for a representative AHU supply fan shows in Fig. 2 that fan power versus fan flow fits a quadratic model.

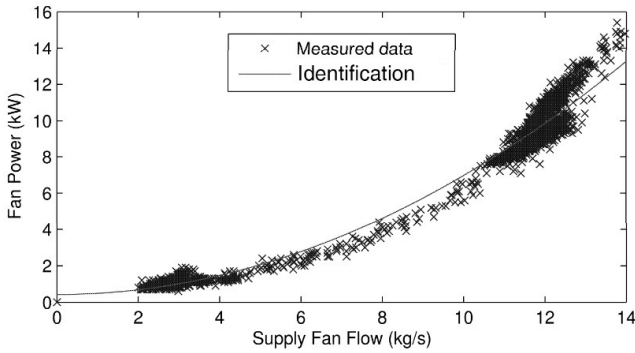


Fig. 2. Recorded data from an AHU supply fan in the Bancroft Library at University of California, Berkeley

With this form of fit and simplification, we have

$$P_f = \kappa_f \left(\sum_{i=1}^n \dot{m}_{zi} \right)^2, \quad (8)$$

where κ_f is a parameter that captures both the fan efficiency and the duct pressure losses.

We introduce several parameters to reflect utility pricing. The cost in currency per unit energy content is denoted r_e for electricity and r_h for heating fuel (typically gas, or steam from a central plant). These costs may vary in time, especially for electricity, to reflect time-of-use or dynamic utility pricing. We assume time variation of utility rates occurs in a zero-order hold manner at sample rate Δt .

We also incorporate a feature of some utility structures wherein peak electric power use is penalized. Some utilities only implement this peak-use charge during certain hours of the day, so we express this feature by defining a windowing function $\phi(t)$. The value of $\phi(t)$ equals the given cost per unit peak power during restricted time intervals, and zero elsewhere.

Assuming cold water is produced immediately on demand by electric chillers, and likewise for hot water by a fuel-powered boiler or district steam heat exchanger, the total utility cost from time t to time $t + N\Delta t$ is

$$J = \int_t^{t+N\Delta t} (r_e P_f + r_e P_c + r_h P_h) d\tau + \max_{\tau \in [t, t+N\Delta t]} (\phi(\tau)(P_f + P_c)), \quad (9)$$

where N is the prediction horizon length.

Note that the cooling coil power (7) depends on the cooling coil inlet air temperature T_m , which depends (5) on return air temperature T_r , which depends (4) on zone temperatures T_{zi} , so this utility cost integral is a function of the state variables and all of the control inputs. The integral in (9) is approximated numerically according to the trapezoidal discretization for consistency with the discrete state dynamics (3).

2.4 Constraints

The system states and control inputs are subject to the following constraints due to control requirements and actuator limits.

- C1: $\underline{T}_{zi} \leq T_{zi} \leq \bar{T}_{zi} \forall i \in \{1, \dots, n\}$, comfort range.
- C2: $\underline{\dot{m}}_{zi} \leq \dot{m}_{zi} \leq \bar{\dot{m}}_{zi} \forall i \in \{1, \dots, n\}$, minimum ventilation requirement and maximum VAV box capacity.
- C3: $T_{si} \geq T_c \forall i \in \{1, \dots, n\}$, heating coils can only increase temperature.
- C4: $T_{si} \leq \bar{T}_h \forall i \in \{1, \dots, n\}$, heating coil capacity.
- C5: $T_c \leq T_m$, cooling coil can only decrease temperature.
- C6: $T_c \geq \underline{T}_c$, cooling coil capacity.
- C7: $0 \leq d_r \leq \bar{d}_r$, minimum is no recirculation, maximum set by required fresh air for indoor air quality.

In all of the above constraints, limit values denoted by \star and $\bar{\star}$ may be time-varying and are assumed to be known in advance. These constraints must hold at all times t .

Due to the trapezoidal discretization of the system dynamics, we approximate the time variation of the zone temperatures T_{zi} as piecewise linear between discrete time steps t and $t + \Delta t$. We treat the disturbance inputs \dot{Q}_{offset} and T_{oa} the same way, as continuous piecewise linear functions of time. All of the control inputs are zero-order held so they are discontinuous piecewise constant functions of time. Due to (5), cooling coil inlet temperature T_m is therefore discontinuous and piecewise linear in time. To enforce constraint C5 at all times we introduce the following additional variables

$$\tilde{T}_r = \frac{\sum_{i=1}^n (\dot{m}_{zi} T_{zi}^+)}{\sum_{i=1}^n \dot{m}_{zi}}, \quad (10)$$

$$\tilde{T}_m = (1 - d_r) T_{oa}^+ + d_r \tilde{T}_r, \quad (11)$$

$$\tilde{P}_c = \frac{c_p}{\eta_c} \left(\sum_{i=1}^n \dot{m}_{zi} \right) (\tilde{T}_m - T_c), \quad (12)$$

and corresponding constraint

$$\text{C8: } T_c \leq \tilde{T}_m.$$

Multiply both sides of (4) and (10) by $\sum \dot{m}_{zi}$:

$$\left(\sum_{i=1}^n \dot{m}_{zi} \right) T_r = \sum_{i=1}^n (\dot{m}_{zi} T_{zi}), \quad (13)$$

$$\left(\sum_{i=1}^n \dot{m}_{zi} \right) \tilde{T}_r = \sum_{i=1}^n (\dot{m}_{zi} T_{zi}^+). \quad (14)$$

Note that when $\sum \dot{m}_{zi} = 0$, the HVAC system consumes no energy and the system dynamics evolve only according to thermal loads \dot{Q}_i . In that case the return temperature is not important, so enforcing (13) and (14) is equivalent for our purposes to enforcing (4) and (10).

3. OPTIMAL CONTROL FORMULATION

Let $x_{k|t}$ be the value of the state vector T_z at time $t + k\Delta t$ predicted at time t , $u_{k|t}$ be the value of the combined control input vector $[\dot{m}'_z, T'_s, d_r, T'_c]'$ at time $t + k\Delta t$ predicted at time t , and $v_{k|t}$ be the value of the vector of auxiliary variables $[T_r, T_m, P_h, P_c, P_f, \tilde{T}_r, \tilde{T}_m, \tilde{P}_c]'$ at time $t + k\Delta t$ predicted at time t .

The system dynamics (3), and auxiliary-variable equalities (5)-(8) and (11)-(14) can all be expressed in the following bilinear form:

$$\frac{1}{2} \begin{bmatrix} x_{k|t} \\ u_{k|t} \\ v_{k|t} \\ x_{k+1|t} \end{bmatrix}' C_{j,k|t} \begin{bmatrix} x_{k|t} \\ u_{k|t} \\ v_{k|t} \\ x_{k+1|t} \end{bmatrix} + d'_{j,k|t} \begin{bmatrix} x_{k|t} \\ u_{k|t} \\ v_{k|t} \\ x_{k+1|t} \end{bmatrix} + e_{j,k|t} = 0 \quad \forall j \in \{1, \dots, n + n_v\}, \forall k \in \{0, \dots, N-1\} \quad (15)$$

The matrices $C_{j,k|t}$ are symmetric and indefinite, and $n_v = 8$ is the number of auxiliary variables.

We cannot influence whether the given initial conditions $x_{0|t} = T_z(t)$ satisfy the comfort constraints C1, so we compactly represent the constraints C1-C8 as

$$A_{k|t} \begin{bmatrix} u_{k|t} \\ v_{k|t} \\ x_{k+1|t} \end{bmatrix} \leq b_{k|t}, \quad \forall k \in \{0, \dots, N-1\}. \quad (16)$$

Let $r_{k|t}$ be the value of the vector $[0, 0, r_h, \frac{r_c}{2}, r_e, 0, 0, \frac{r_c}{2}]'$ at time $t + k\Delta t$. Then by the trapezoidal discretization,

$$\int_t^{t+N\Delta t} (r_e P_f + r_e P_c + r_h P_h) d\tau = \sum_{k=0}^{N-1} \Delta t r'_{k|t} v_{k|t}. \quad (17)$$

Lastly, we introduce an epigraph variable β_t to represent the peak power term in (9).

Let $F_{k|t} = \begin{bmatrix} 0 & 0 & 0 & \phi(\tau) & \phi(\tau) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \phi(\tau) & 0 & 0 & \phi(\tau) \end{bmatrix}$ evaluated at $\tau = t + k\Delta t$. Then we have the equivalence that

$$\max_{\tau \in [t, t+N\Delta t]} (\phi(\tau)(P_f + P_c)) = \min \beta_t \text{ s.t. } F_{k|t} v_{k|t} \leq \begin{bmatrix} \beta_t \\ \beta_t \end{bmatrix}, \quad \forall k \in \{0, \dots, N-1\}. \quad (18)$$

The constrained finite time optimal control (CFTOC) problem at time t is then

$$\begin{aligned} \min_{\mathbf{x}_t, \mathbf{U}_t, \mathbf{V}_t, \beta_t} \quad & \beta_t + \sum_{k=0}^{N-1} \Delta t r'_{k|t} v_{k|t} \\ \text{s.t.} \quad & \forall j \in \{1, \dots, n + n_v\}, \forall k \in \{0, \dots, N-1\}, \\ & \frac{1}{2} \begin{bmatrix} x_{k|t} \\ u_{k|t} \\ v_{k|t} \\ x_{k+1|t} \end{bmatrix}' C_{j,k|t} \begin{bmatrix} x_{k|t} \\ u_{k|t} \\ v_{k|t} \\ x_{k+1|t} \end{bmatrix} + d'_{j,k|t} \begin{bmatrix} x_{k|t} \\ u_{k|t} \\ v_{k|t} \\ x_{k+1|t} \end{bmatrix} + e_{j,k|t} = 0 \\ & A_{k|t} \begin{bmatrix} u_{k|t} \\ v_{k|t} \\ x_{k+1|t} \end{bmatrix} \leq b_{k|t}, \quad F_{k|t} v_{k|t} \leq \begin{bmatrix} 1 \\ 1 \end{bmatrix} \beta_t, \quad x_{0|t} = T_z(t) \end{aligned} \quad (19)$$

where $\mathbf{U}_t = \{u_{0|t}, \dots, u_{N-1|t}\}$ is the set of predicted control inputs at time t , $\mathbf{X}_t = \{x_{1|t}, \dots, x_{N|t}\}$ is the set of predicted system states at time t , starting from initial state $x_{0|t} = T_z(t)$ and applying the input sequence \mathbf{U}_t to the system model (3), and $\mathbf{V}_t = \{v_{0|t}, \dots, v_{N-1|t}\}$ is the corresponding set of predicted auxiliary variables at t .

Let the optimal solution of problem (19) have input sequence denoted by $\mathbf{U}_t^* = \{u_{0|t}^*, \dots, u_{N-1|t}^*\}$. Then, the first step of \mathbf{U}_t^* is input to the system, $u(t) = u_{0|t}^*$. The optimization (19) is repeated at time $t + \Delta t$, with the updated new state

$$x_{0|t+\Delta t} = T_z(t + \Delta t)$$

yielding a *moving or receding horizon control* strategy. The values of $b_{k|t+\Delta t}$, $d_{j,k|t+\Delta t}$, $e_{j,k|t+\Delta t}$, $F_{k|t+\Delta t}$, and $r_{k|t+\Delta t}$ may change relative to the previous time step due to time-varying constraints, predictions, and utility costs.

3.1 Sequential Quadratic Programming Approach

The optimization problem (19) is nonconvex due to the nonlinear equality constraints (15). The method of sequential quadratic programming (SQP) is applied here by linearizing (15) and adding a convex quadratic term to the cost that approximates the Hessian of the Lagrangian function, see Han (1977); Nocedal and Wright (2006).

Let $\lambda_{j,k|t}$ be the Lagrange multiplier corresponding to equality constraint j of step k in problem (19) at time t . The Hessian of the Lagrangian of problem (19) is then

$$\begin{aligned} & \sum_{j=1}^{n+n_v} \left(\begin{bmatrix} 0_{n \times (n+n_u+n_v)N} \\ I_{(n+n_u+n_v)N} \end{bmatrix}' \begin{bmatrix} \lambda_{j,0|t} C_{j,0|t} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0_{n \times (n+n_u+n_v)N} \\ I_{(n+n_u+n_v)N} \end{bmatrix} \right. \\ & \left. + \sum_{k=1}^{N-1} \begin{bmatrix} 0_{(n+n_u+n_v)k-n} & 0 & 0 \\ 0 & \lambda_{j,k|t} C_{j,k|t} & 0 \\ 0 & 0 & 0_{(n+n_u+n_v)(N-k-1)} \end{bmatrix} \right) \quad (20) \end{aligned}$$

where $n_u = 2n + 2$ is the number of control inputs. The first n rows and first n columns of $C_{j,0|t}$ for constraints at time step $k = 0$ are removed because the initial state $x_{0|t}$ is a known quantity. Since each matrix $C_{j,k|t}$ is indefinite, this exact Hessian of the Lagrangian is also indefinite.

We can efficiently calculate the smallest eigenvalue and corresponding eigenvector of this Hessian using an Arnoldi iteration, see Lehoucq et al. (1998). We construct a positive semidefinite approximation of the Hessian by adding a multiple of the identity to (20). Using the available second derivative information in this manner was observed to reduce the number of iterations required for convergence relative to the quasi-Newton approaches of the commercial SQP routines we tested.

4. SIMULATION RESULTS

We present simulation results for the following cases:

- (1) Nominal case, r_e uses “low” value from Table 2 and $\phi(t) = 0$ at all times.
- (2) Modified electric rate schedule: r_e uses “high” value between 12 noon and 4:30 PM and the “low” value at all other times. No peak power charge, $\phi(t) = 0$.
- (3) Peak power penalty: $\phi(t) = 1$ \$/kW and r_e uses “low” rate at all times.

Table 2. Parameter values used

Parameter	Value	Units
n	5	zones
N	48	steps
Δt	1800	s
c_p	1	kJ/(kg·K)
$(mc)_i$	1000	kJ/K
R	0	kW/K
η_h	0.9	dimensionless
η_c	4	dimensionless
κ_f	0.065	kW·s ² /kg ²
r_e	$\begin{cases} \text{"high"} = 1.5 \cdot 10^{-4} \\ \text{"low"} = 3 \cdot 10^{-5} \end{cases}$	\$/kJ
r_h	$5 \cdot 10^{-6}$	\$/kJ
T_0	18	°C
\underline{T}_{zi}	$\begin{cases} 6:30 \text{ AM to } 6:30 \text{ PM} = 21 \\ 7 \text{ PM to } 6 \text{ AM} = 12 \end{cases}$	°C
\bar{T}_{zi}	$\begin{cases} 6:30 \text{ AM to } 6:30 \text{ PM} = 24 \\ 7 \text{ PM to } 6 \text{ AM} = 32 \end{cases}$	°C
$\underline{\dot{m}}_{zi}$	$\begin{cases} 6:30 \text{ AM to } 6:30 \text{ PM} = 0.025 \\ 7 \text{ PM to } 6 \text{ AM} = 0 \end{cases}$	kg/s
$\bar{\dot{m}}_{zi}$	1.5	kg/s
\underline{T}_c	5	°C
\bar{T}_h	40	°C
\bar{d}_r	0.9	dimensionless

For ambient temperature T_{oa} , we use a sinusoid with period 1 day, minimum value 10 °C at time 1:30 AM, and maximum value 30 °C at 1:30 PM. Zone thermal loads \dot{Q}_i are set to the time-varying profiles shown in Fig. 3.

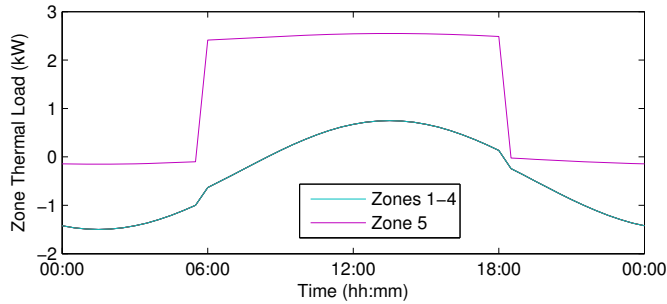


Fig. 3. Zone thermal loads \dot{Q}_i

The SQP iterations were computed using the solver BPPMD, see Mészáros (1996). The results of case 1 are shown in Fig. 4 and 5, case 2 is shown in Fig. 6, and case 3 is shown in Fig. 7. More extensive parametric simulations were executed but are not included here due to space constraints. Those results can be found in a technical report available at <http://www.mpc.berkeley.edu/>.

We observe several interesting behaviors for the nominal case 1 in Fig. 4 and 5. Before the occupied period begins at 6:30 AM, zone 5 is set to zero flow and zones 1-4 have a small flow rate at the maximum heating temperature to counteract the cooling loads (negative \dot{Q}_i). All zones preheat to satisfy the tighter occupied temperature constraints before the occupied hours begin. From 6:30 AM until 9 AM, zone 5 is in cooling mode but the other zones are in heating. This period is an economizer mode condition: a mix of outside air maintains the mixed temperature close to \underline{T}_{zi} , while zone 5 satisfies its cooling demand with a very large flow rate. After 9 AM the ambient temperature is warmer than the return temperature so the AHU dampers return to maximum recirculation. The cooling

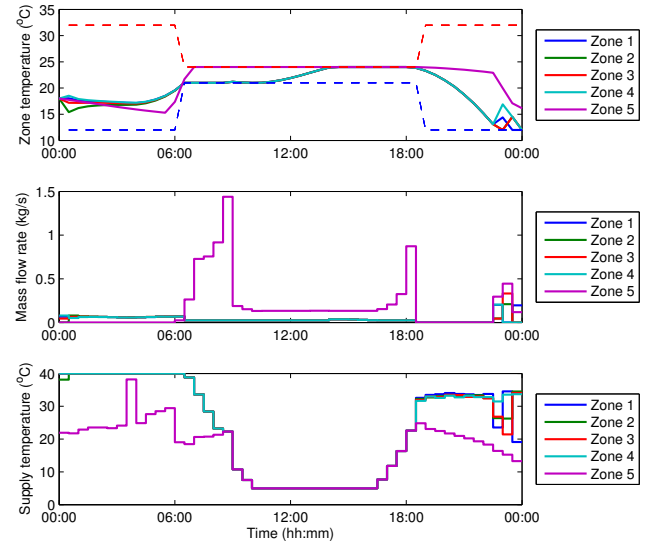


Fig. 4. Case 1 zone results. Shown dashed in the first plot are \underline{T}_{zi} and \bar{T}_{zi} .

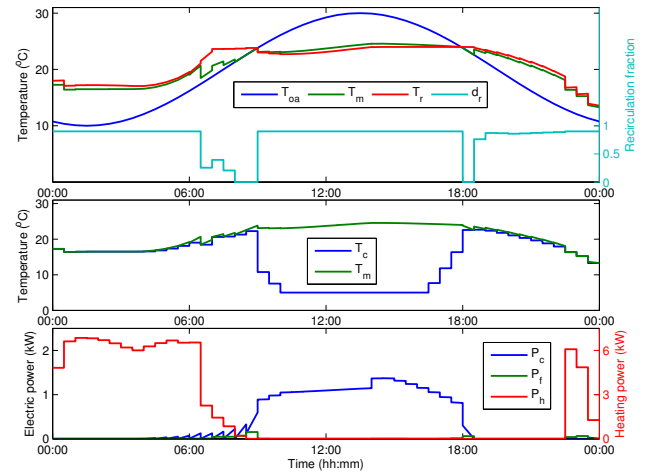


Fig. 5. Case 1 AHU results.

coil activates at this time, reaching its lowest setpoint by 10AM. Zones 1-4 begin transitioning to cooling mode here. Immediately before the end of the occupied period at 6:30 PM, we see a cooling coil supply temperature reset behavior. The load prediction is much lower after 6:30 PM, so the cooling coil begins increasing its setpoint early, trading lower cooling power for higher fan power (the flow to zone 5 must increase to keep it cooled using warmer supply air). The erratic one-at-a-time heating of zones 1-4 after 10:30 PM appears to be a consequence of the return temperature dependence on mass flows. When only one zone is heated with a large mass flow (others at low flow), the return temperature is influenced most by the high-flow zone. Increased return temperature reduces the required heating coil energy for the next zones to be heated.

Both case 2 in Fig. 6 and case 3 in Fig. 7 demonstrate precooling of zone 5 and lengthened cooling of zones 1-4, but with different timing and intent. In case 2, $\phi(t) = 0$ but the electric rate r_e has a higher value between 12 noon and 4:30 PM. So precooling is only performed immediately before noon, with a corresponding spike in cooling power,

so that less cooling energy is used between 12 noon and 4:30 PM. In case 3, $\phi(t) = 1$ so load-shifting is used to minimize peak power. Zone 5 is pre-cooled beginning earlier in the morning, increasing cooling power at a time when it would otherwise be low and shifting electric power use away from the times it would normally be highest.

Our computational results show that in response to either time-varying electric rates (r_e , case 2) or peak power penalties ($\phi(t)$, case 3), this optimization-based controller does not use appreciably more total energy. We are not showing the combination of r_e high and $\phi(t) = 1$ here, but the results are very similar to case 3. Imposing a penalty on peak power rules out the type of short-duration precooling seen in case 2.

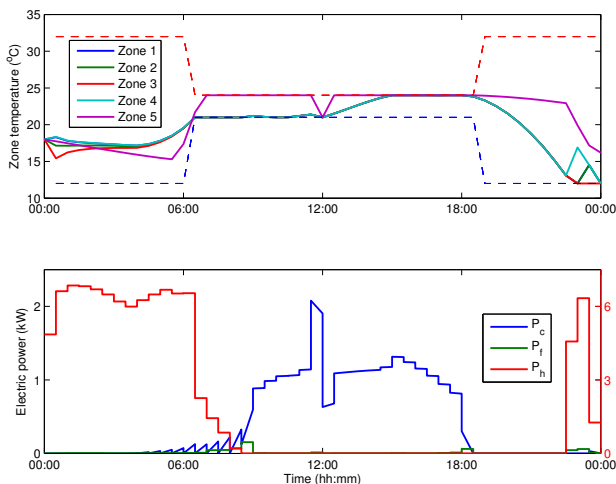


Fig. 6. Case 2 overview. Note the precooling and spike in cooling power immediately before noon.

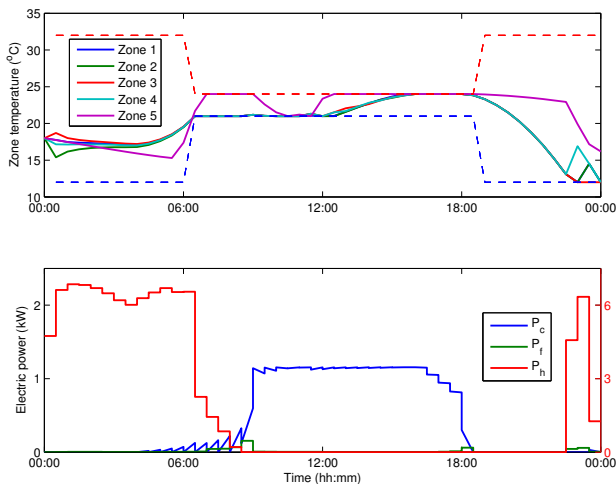


Fig. 7. Case 3 overview. Note the timing of the precooling and the intentional plateau in cooling power.

5. CONCLUSIONS

This MPC algorithm has shown encouraging results in a few interesting cases for a sample problem. The control performance, entirely from an optimization origin, exhibits aspects of heuristic HVAC control such as economizer

control, supply temperature reset, demand response, pre-cooling, and load-shifting in a coordinated manner. The computational time for each of the above cases was less than one minute, faster than the time scales of a HVAC system. We are working toward real-time receding-horizon implementation of this algorithm to experimentally control a building. Future work is necessary in the areas of system identification, model validation, and thermal load prediction. We also plan on investigating robust MPC for this system, wherein we account for uncertainty in future thermal load values and the effects of model mismatch. Lastly, the SQP algorithm only guarantees local solutions to non-convex nonlinear optimization problems, so we are investigating branch-and-bound extensions of this approach for global search algorithms.

ACKNOWLEDGEMENTS

This work was supported by the Department of Defense under the award W912HQ-09-C-0056. We thank Allan Daly and Yudong Ma for discussions regarding HVAC modeling and optimization.

REFERENCES

- Borrelli, F. (2003). *Constrained Optimal Control of Linear and Hybrid Systems*, volume 290. Springer-Verlag.
- Han, S. (1977). A globally convergent method for nonlinear programming. *Journal of Optimization Theory and Applications*, 22(3), 297 – 309.
- Henze, G., Felsmann, C., and Knabe, G. (2004). Evaluation of optimal control for active and passive building thermal storage. *International Journal of Thermal Sciences*, 43(2), 173 – 183.
- Lehoucq, R., Sorensen, D., and Yang, C. (1998). *ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*. SIAM Publications.
- Ma, Y., Borrelli, F., Hencsey, B., Coffey, B., Bengae, S., and Haves, P. (2010). Model predictive control for the operation of building cooling systems. In *2010 American Control Conference*, 5106 – 5111.
- Mayne, D., Rawlings, J., Rao, C., and Sokaert, P. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, 36(6), 789 – 814.
- McQuade, J. (2009). A systems approach to high performance buildings. Technical report, United Technologies Corporation. <http://gop.science.house.gov/Media/hearings/energy09/april28/mcquade.pdf>.
- Mészáros, C. (1996). Fast cholesky factorization for interior point methods of linear programming. *Computers & Mathematics with Applications*, 31(4-5), 49 – 54.
- Nocedal, J. and Wright, S.J. (2006). *Numerical Optimization*, chapter 18. Springer-Verlag.
- Oldewurtel, F., Parisio, A., Jones, C., Morari, M., Gyalistras, D., Gwerder, M., Stauch, V., Lehmann, B., and Wirth, K. (2010). Energy efficient building climate control using stochastic model predictive control and weather predictions. In *2010 American Control Conference*, 5100 – 5105.
- Qin, S. and Badgwell, T. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, 11, 733 – 764.