Predictive Control Methods to Improve Energy Efficiency and Reduce Demand in Buildings

Tim Salsbury [*], Prashant Mhaskar [†], and S. Joe Qin [‡]

Abstract

This paper presents an overview of results and future challenges on temperature control and cost optimization in building energy systems. Control and economic optimization issues are discussed and illustrated through sophisticated simulation examples. The paper concludes with effective results from model predictive control solutions and identification of important directions for future work.

1. Introduction

The need for control in buildings usually resides in the mechanical and electrical systems that are installed to maintain a comfortable and safe indoor environment. A wide range of these systems can be found in buildings including heating, ventilating, air-conditioning (HVAC), lighting, security, elevators, escalators, fire detection and abatement. All these systems use energy and produce useful work as output. In the case of HVAC, energy is used to maintain temperature, humidity, and air quality at levels in accordance with the building purpose. The energy process is illustrated in Figure 1.

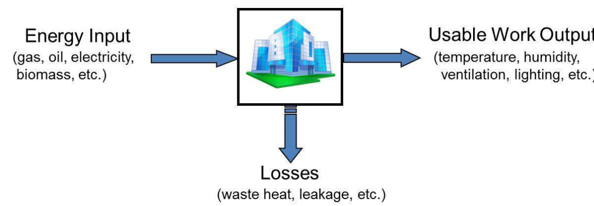

Figure 1. Energy process in a building

As with any generic process control problem, the objective is to satisfy the output requirements in the most efficient way and hence with the least amount of input (energy). Because energy is a cost, control in buildings, as in most other applications, can be translated to an economic optimization problem. In simple mathematical terms, the problem can be stated as the minimization of the integral of the energy usage subject to constraints on the measured variables (see Section 2 for an example). The real problem, however, is more complex because the constraints are not linear and the true goal is to minimize the energy cost (not just the energy usage), which is also nonlinear and time varying.

Although economic optimization might be the overall goal, it is very difficult to address this directly. The reason for this stems from the many constraints and intermediate objectives that are the result of buildings being large complex processes. The design process for control in buildings has therefore traditionally been from the bottom up [1]. What this means is that control logic is designed around each constraint and internal objective and then coalesced to form the overall control strategy. The rest of this section describes control problems at these lower levels in buildings and for the sake of brevity, focuses on results in the literature specific to the building control problem.

Single variable regulation: This is the most basic control task and involves maintaining a variable to a set-point by manipulating a device. An example would be maintaining 72 degree temperature in a conference

[*] Johnson Controls Inc., 507 E. Michigan Street, Milwaukee, Wisconsin, United States 53202

[†] Department of Chemical Engineering, McMaster University, Hamilton, Ontario, Canada L8S 4L7

[‡] Mork Family Department of Chemical Engineering and Materials Science University of Southern California 925 Bloom Walk, HED 211 Los Angeles, CA 90089-1211

room by moving a damper to regulate the amount of cooled air entering the room. Feedback control is the most common solution for these problems and PID is the most common algorithm that is used. There are several issues that arise in buildings that make the single variable feedback problem more challenging. Examples are: hysteresis, stiction, static non-linearity, time variance, capacity problems, dead-zones, quantization, time-delays. Many of these problems are more prevalent in buildings compared to other control applications because the low cost nature of the industry leads to unreliable and inaccurate components [2].

Multi-variable control: Many of the single variable control problems in buildings are part of broader multi-variable control problems. The approach of treating each problem in isolation is common though and the design therefore becomes that of a decentralized multi-loop strategy. Cascaded control is used in buildings typically via multiple PID controllers. A specific example of a multi-variable control problem is the regulation of both humidity and temperature. Different devices are used to control these variables but a change in one affects the other. Most often, interactions of these types are ignored and control performance is sub-optimal [3].

Scheduling: There are many scheduling problems in buildings used to determine when to turn on and off equipment to satisfy load demands. Time-based scheduling is common where algorithms are used to determine when to turn on a device in order to reach desired conditions by a certain time. A specific example of this is "optimal start", which is used to turn on heating or cooling devices in buildings so that temperatures in zones are at setpoints when the occupancy time is reached [4]. Other types of scheduling problems involve determining how to operate multiple devices in series or in parallel. Devices that are in series are most commonly handled using split range control or state-based logic [5]. An example is moving from heating to cooling by splitting the output of a PID controller so that it maps on to two device inputs. Examples of parallel systems would be equivalent devices that serve different parts of a building. For example, a building might have multiple air-ventilation systems for different sections of a building. Coordination of these systems is rare, but algorithms are sometimes deployed to take account of start-up and shut-down events in order to minimize load spikes.

Constraint imposition: There are a myriad of safety features built into all types of building systems. These may be hardware or software based and the goal is to prevent damage to the equipment by constraining the range of operation. More expensive pieces of equipment such as chiller systems will have more safety measures to prevent damage to components. The state of the art in the buildings area is that the safety and operational constraints are usually implemented in parallel to the control logic. Conflicts then often arise between the two sets of logic and safety-trips are common.

Resource management: A typical building will have significant redundancy meaning that demands can be met by using different combinations of equipment. An example for an HVAC system is when both the ventilation rate and the temperature of the air delivered to spaces can be varied to achieve a desired heat transfer rate. The control logic used to determine these trade-offs is usually based on experience that has been built up over years. Mode changes and so-called "supervisory" control logic is used to manage this kind of redundancy but the design is once again bottom up and it is rare to find system level optimization built into the design of these strategies. Some of the scheduling problems mentioned earlier could also be grouped in this resource management category and there is an obvious similarity between these two problem areas.

Fault tolerance. In addition to faults resulting from equipment failure, many building system components are non-ideal. Examples were given earlier of non-linearity that is prevalent in building systems and control logic has to be designed to, at least, be robust to these expected operational deficiencies. Adaptive control at the PID level has been used in buildings to address static nonlinearity and time-variant behavior [6]. Also gain scheduling and static gain inverters can be found. Filtering including smoothing and spike removal is also used in control logic. The system redundancy mentioned earlier coupled with the decentralized control design also means that there is a certain amount of fault tolerance built in to most building systems. This means that setpoints can be met under most conditions even when some items of equipment have failed. This is useful from an operational point of view but can lead to energy penalties and make fault detection difficult because faults are masked by the redundancy [7, 8].

Opportunities: Modern control systems in buildings are implemented on a digital communication network that has different layers with the higher levels being IP-based over Ethernet. The higher level control devices are equivalent to PCs and usually connect out to the Internet. Traditionally, the building control system has been separate from the rest of a building's business and telecommunication IT infrastructure. However, because both networks are essentially the same, there is a convergence underway, particularly as building owners see the economic benefits of combining parts of the networks rather than installing parallel duplicates.

As network communication components have become commoditized this has allowed for an increasing number of devices in a building to be connected to a network. This interconnectedness is creating new opportunities for control design as data from many sources can be accessed at any point on the network. However, the evolution of the networking infrastructure has outpaced the development of control algorithms that can take advantage of these new capabilities. The design, as mentioned earlier, is still decentralized multi-loop that is built from the bottom up. Opportunities therefore exist for broadening the control design to encompass groups of systems so that system-level optimization can take place at a higher level than is currently done.

If the ultimate goal of control in buildings is one of (constrained) economic optimization, this can be achieved in one of two ways. Either optimization is carried out for the entire building whereby total energy costs are minimized, or costs at several sub-levels are minimized independently. Only in theory would be possible to find the (Pareto) optimal point by optimizing a building in its entirety, because this would demand that all information be available with certainty to a central location, known as the economic calculation problem in economics [9]. This is not possible in practice and any efforts to try to model an entire building will be hampered with significant inaccuracy and uncertainty. The theoretical optimal solution will therefore be unrealizable in practice and may also lead to building-wide problems when models, predictions, or measurements break down.

The alternative approach of breaking the optimization problem into smaller pieces is the only practically realizable option for a large building (see Figure 2). Although this approach may lead to a sub-optimal point, the approach has several important advantages. First, inevitable uncertainty and modeling errors will only affect localized parts of the building thus resulting in risk diversification and enhanced robustness. Second, the approach is an incremental change over current state of the art thereby making it more viable for industrial adoption. Moving up the system hierarchy to the point where price signals are available is nevertheless a radical improvement over current practice that could lead to significant economic benefits. Finally, once economic optimization is implemented performance could be further enhanced by incorporating cooperative strategies that take advantage of the potential to share information between each of the independent optimizing agents (e.g., [10]).

The two research efforts presented in this paper adopt this more system-centric approach to control design by applying model predictive control (MPC) methods to building applications. Both of the MPC control strategies that are described are multi-variable in nature and thus take advantage of the ability of building control systems to consolidate measurements from several sources. The proposed control strategies also incorporate economic optimization as well as setpoint regulation. This is an important improvement over the current state of the art that is made possible by employing control logic at a high enough level where an economic (price) signal is measurable. The rest of the paper is organized as follows: In Section 2 results on advanced higher level control of buildings are presented followed by an overview of existing results on the problem of control of cooling units in Section 2 as well as illustrative results. Next, in Section 3, load shifting techniques for energy optimization are discussed, followed by a discussion on future directions in Section 4. Finally, we present the conclusions in Section 5.

2. Cooling Plant Control

In this section, we review results on control of cooling units. We first present an overview of control results on cooling units (including experimental as well as simulation based studies) then focus on cooling units using a Vapor Compression Cycle (VCC) as the cooling mechanism. We then discuss classical control approaches for VCC, followed by model-based approaches for control of VCC units, and present some illustrative results.

2.1. System Description and Control Relevant Issues

In the recent past there have been a significant number of simulation and experimental-based studies on the modeling and control of cooling systems including refrigeration systems for food preservation, roof top cooling units, and chillers (see, e.g., [11, 12, 13, 14, 15, 16, 17] and the references therein). For the sake of highlighting the key control relevant issues, we focus here on VCC, a representative (and quite commonly used) cooling mechanism in HVAC systems, for which excellent models (validated using experimental data) covering various regimes of operations of vapor compression cycles have been reported in a series of papers ([18, 19, 20, 21] and [22] for an excellent review).

A vapor compression cycle (VCC) refers to a type of thermodynamic machinery that transfers heat using a compressible fluid referred to as the refrigerant. The most common realization of a VCC consists of four components: a compressor, condenser, expansion valve, and an evaporator. In a VCC unit, the refrigerant enters the compressor as a superheated vapor and is com-

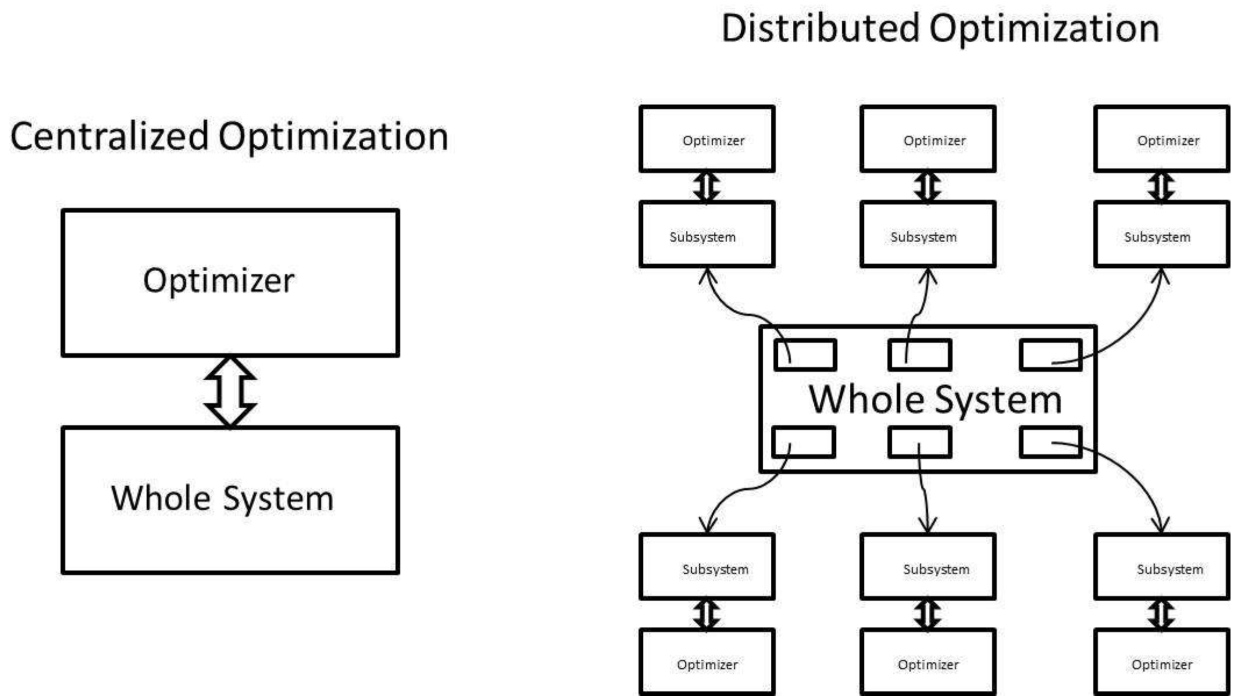

Figure 2. Centralized versus distributed optimization

pressed to a higher pressure, resulting in superheated vapor. From the compressor, the superheated refrigerant vapor enters a condenser (typically placed outdoors), condensing to a sub-cooled liquid at the condenser exit as a fan blows the ambient air over the condenser. The high pressure sub-cooled liquid then flows into an expansion valve which decreases the pressure and temperature of the refrigerant, causing a liquid-vapor mixture to form. Then, the two-phase refrigerant mixture enters an evaporator that is exposed to the environment to be cooled. The environment temperature is above the temperature of the refrigerant, resulting in the evaporation and subsequent heating of the refrigerant to a superheated vapor at the evaporator exit. The cooling medium (air or water), in turn, is cooled and available to be distributed for cooling. The superheated vapor from the evaporator exit then flows back into the compressor, completing the cycle.

The control objectives are typically defined in terms of degrees of superheat in the refrigerant at the evaporator exit and the air temperature (or the temperature of the cooling medium) at the evaporator exit (the supply air temperature). Ensuring that superheated refrigerant exits the evaporator is of utmost importance in preventing physical damage in the VCC, as liquid refrigerant can damage the mechanical components used in the compressor. At the same time, maximization of the two-phase region within the evaporator maximizes the energy efficiency of the cycle resulting in a tradeoff between optimality and safety. The manipulated variables include the compressor speed, the air flow rates across the condenser and the evaporator and the expansion valve opening. The compressor is the largest energy consumer, with the energy consumption typically being a polynomial, increasing function of the compressor RPM.

2.2. VCC Control

In this section, we review applications of classical control approaches as well as more recent results on model-based control approaches of VCC.

Classical control approaches: When using single input single output control approaches (classical PI/PID or various versions of it), the success of the control structure relies to a large extent on the choice of the control-loop pairing. The performance of the system is limited by the inherent nonlinearity and multivariable nature of the problem. Typical control loop pairings include superheat-valve opening and cooling load (captured through evaporator pressure or supply air temperature)-compressor RPM. Recent results that utilize classical control structures, albeit using tools such as static decouplers, adaptive PID designs as well as

cascaded control structures [22, 23, 24] demonstrate better performance over simple PI loops as well as point to the possibility of improved performance achievable using advanced control structures.
Model-based control approaches: One example of a Model-based control design is the use of a reduced order model, developed from a first principles nonlinear model, for the purpose of control design [22]. In [24] a model-based predictive controller is presented, where a discrete time input-disturbance-output model is identified from 'data' obtained from the detailed nonlinear model used for the purpose of simulations. In particular, an autoregressive with exogenous terms (ARX) model of the form shown below is utilized:

$$y(k) = \sum_{i=1}^{n_y} \mathbf{A}_i y(k-i) + \sum_{i=1}^{n_u} \mathbf{B}_i u(k-i) + \sum_{i=1}^{n_d} \mathbf{C}_i d(k-i) \tag{1}$$

where $y(k)$ and $u(k)$ are the process output and input vectors at sampling instant k (respectively), $d(k)$ is a vector of measurable disturbances, $\mathbf{A}_i$, $\mathbf{B}_i$, and $\mathbf{C}_i$ are model coefficient matrices (that are estimated using least-squares regression), and n_y, n_u, and n_d denote the (maximum) number of time lags in the outputs, inputs, and disturbances (respectively) and define the order of the model. For specific outputs, inputs, or disturbances which do not require the maximum number of lags, the appropriate elements in the coefficient matrices can be set to zero. Note that for the VCC, the mixed air temperature (the inlet temperature of the air flowing over the evaporator) and the ambient temperature are natural (and measured) disturbance variables that can be incorporated as the disturbance variables for improved modeling.
Model Predictive Control (MPC): The key components of a model predictive controller (see, [25] for an excellent review) are 1) a model (linear or nonlinear) that allows prediction of the process states for candidate input trajectories and 2) description of the objective function/constraints that reflect the desired (or as limited by physical constraints) behavior of the process. Applications of MPC are become increasingly prevalent due to its ability to handle the multivariable/nonlinear nature of the dynamics, constraints and optimality in an integrated fashion, and has been the subject of several research studies (see e.g., [26, 27, 28, 29] for recent results that establish stability guarantees from well characterized operating regions while handling input [26] and state constraints [27] as well as uncertainty [28], while enhancing the stability region achievable using existing control approaches [29]).

Recently, model predictive control formulations have been implemented on VCC test beds. In [30, 31] a linear model based (identified from test data) predictive controller is implemented on a multi-evaporator compression cooling cycle, where the energy efficiency results from the ability of the MPC to keep the superheat at as low values as possible. In [32] a nonlinear model predictive controller (based on a reduced order, first principle, nonlinear model) is implemented on a test bed and performance improvement demonstrated. In [33], a gain-scheduling approach is utilized to handle the nonlinearity.

To illustrate the key idea behind a typical MPC implementation, consider a predictive control design [24] where the inputs to the VCC at sampling instant i are computed by solving the following constrained optimization problem:

$$\min_{u_{\mathsf{min}} \le u(k) \le u_{\mathsf{max}}} \sum_{k=1}^{P} \|\hat{y}_2^*(k) - y_{2,\mathsf{SP}}(k)\|_{\mathbf{Q}} + \|u_1\|_r + \|\Delta u\|_{\mathbf{R}}$$

st: $\Delta u_{\mathsf{min}} \le \Delta u(k) \le \Delta u_{\mathsf{max}}$

$Eq.1$

$$\hat{y}^*(k) = \hat{y}(k) + \alpha + \beta(i)$$
$$y_{1,\mathsf{min}} \le \hat{y}_1^*(k) \le y_{1,\mathsf{max}}$$
$$\alpha = k\,[y(0) - \hat{y}(0)]$$
$$\beta_1(i) = \beta_1(i-1) + g_1 \max\{0, [y_{1,\mathsf{min}} - y_1(0)]\}$$
$$+ g_2 \max\{0, [y_1(0) - y_{1,\mathsf{max}}]\}$$
$$\beta_2(i) = \beta_2(i-1) + f\,[y_2(0) - y_{2,\mathsf{SP}}(0)]$$

where the notation, $\|\cdot\|_{\mathbf{Q}}$, refers to the weighted norm, defined by $\|x\|_{\mathbf{Q}} = x^{\mathrm{T}}\mathbf{Q}x$ and Δu denotes a vector in which each element is the difference between successive input moves. The weighting matrices are diagonal and used to trade-off the relative importance of the different control objectives. The plant measurement at the current sampling instant i corresponds to $k = 0$ or $y(0)$.

In this MPC formulation, the control objective of supply air temperature set-point tracking is addressed by penalizing the deviation between the predicted supply-air temperature from its set-point, $y_{2,\mathsf{SP}}(k)$, weighted by $\mathbf{Q}(1,1)$. The predicted superheat is also bounded between $y_{1,\mathsf{min}}$ and $y_{1,\mathsf{max}}$. To reduce the energy consumption associated with the control action, the value of RPM is also penalized using the weight r. The inputs are constrained in a range for which the nonlinear VCC model is known to be valid. In addition to using hard constraints for the input rates, excessive input movements are penalized in the objective function using a move suppression factor with the weighting matrix, $\mathbf{R}$. When tuning the different weighting matrices, the highest importance was initially given to tracking the supply air set-point. Subsequently, the remaining weighting matrices were adjusted appropriately to achieve relatively smooth input behavior.

To achieve offset free performance, a disturbance/bias term is added to the model predictions that is expressed by combining two *constant* terms, α and $\beta(i)$. The first term, α, is the disturbance due to plant-model mismatch at the current sampling instant, multiplied by a tuning parameter, k. Specifically, α is defined as the difference between the predicted outputs at sampling instant i from the *previous* control calculation and the measured outputs at i. The $\beta(i)$ term is the summation of tracking errors up to and including sampling instant i. For the superheat (output 1 or y_1), a non-zero tracking error at i is used only if the current measurement exceeds the minimum or maximum superheat. The β term essentially "persists" and influences the control action until the offset is eliminated. It can be understood as operating the same way as the integral mode in a PI controller. The tuning parameters, g_1, g_2, and f, are used to trade-off the input aggressiveness and the amount of offset (for the parameter values, see [24]). Figure 3 demonstrates the effect of the α and β. In the nominal case (no corrections), there is considerable offset in the supply air temperature. After adding the feedback term to account for plant-model mismatch, this offset is significantly reduced but not eliminated. Zero offset is only achieved after also including the error accumulation term in the formulation.

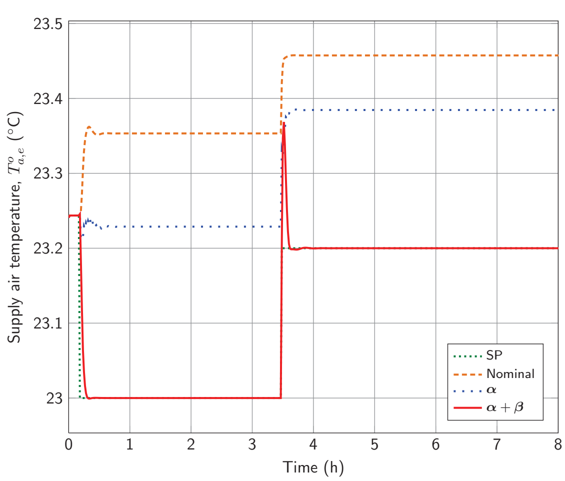

Figure 3. Supply air temperature responses using various combinations of the bias terms in the proposed MPC design

Next, closed-loop simulation results for MPC and PI control are compared. For these simulations, constant disturbances are assumed. That is, the ambient air conditions (temperature and humidity) and the inlet air temperature to the evaporator (the mixed air temperature) are maintained at constant values. Using the results in [34], for the PI loop pairing, the supply air temperature is paired with the compressor RPM while the superheat is paired with the expansion valve opening. The superheat set-point for the PI controller is specified to be 10°C. The PI controllers are initially tuned using the internal model control tuning method and fine-tuned to minimize the integral of absolute error while maintaining relatively smooth input trajectories.

Table 1. Stand-alone VCC closed-loop performance metrics

	Control strategy	
Metric	PI Control	MPC
$\mathsf{ISE}_{\mathsf{SA}}$ (s · °C^2)	837	222
TEC (kJ)	10017	9217

Figure 4 displays the closed-loop VCC responses for the two control strategies (the input profiles are omitted for brevity) and Table 1 summarizes their control performances using the metrics $ISE_{SA} = \Delta t \sum_{i=1}^{K} \left[T^o_{a,e,\mathsf{SP}}(i) - T^o_{a,e}(i) \right]^2$ and the total energy consumption (TEC) over the duration of the simulation (denoted by K simulation steps), for the different set-point step changes. As shown in Figure 4, the proposed MPC design is able to provide better tracking performance of the supply air temperature for the different set-point changes with similar settling times and lower energy consumption. The third supply air set-point change (to approximately 23.7°C) is an infeasible set-point for the VCC cooling capacity, but note that the predictive controller is able to drive the supply air temperature closer to this set-point compared to the PI controller. Note, however, that the infeasibility is merely a result of the model not being valid at low RPM (or as low as required) to provide less cooling.

For the MPC design, the superheat is permitted to "float" between its minimum and maximum value whereas for PI control, the superheat is maintained at the constant safety margin of 10°C. This additional 'degree of freedom' for the predictive controller leads to more accurate tracking and better overall control performance. Note that if the superheat was prescribed to be maintained at a constant value of 10°C for the MPC design as well, the corresponding closed-loop results would be similar to those obtained when using the PI controller. In regards to the energy efficiency, the MPC design required 8% less energy compared to the PI controller. This is a consequence of using higher valve openings and lower RPM values resulting from the multivariable nature of the MPC controller and the ability to allow the superheat to 'float' between acceptable values.

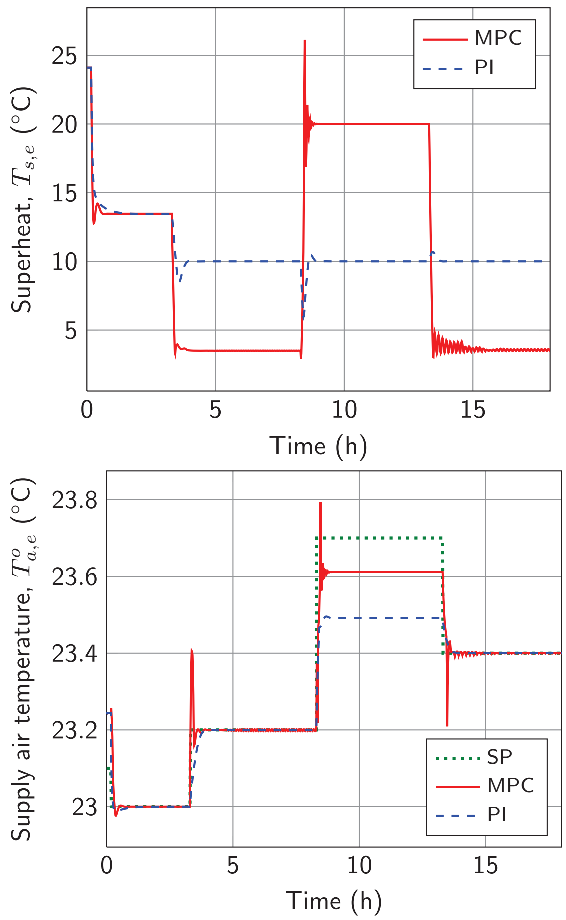

Figure 4. Closed-loop output profiles for the VCC under MPC and PI control.

3. Load Shifting and Demand Reduction via the MPC Framework

In the United States about 70% of electricity is consumed in commercial and residential buildings. To make it worse, the peak demands of building cooling (or in some regions heating) usually occur around the same time period during the day, making the electricity consumption at the peak time (known as *demand*) extremely high relative to the average consumption level. The high peak demand dictates that the power generation capacity from available power plants has to be at least equal to the peak demand, or a blackout would occur. The issues associated with concentrated high demand are twofold. One is that new power plants have to be built to meet the increased demand. The other is that during off-peak hours many power plants have to be shut down or operated at a reduced load, sacrificing power generation efficiency and overall equipment utilization. The situation is worsened by the integration of renewable energy such as wind energy, which generates more during night, which is the off-peak consumption time. If the peak demand can be reduced by properly making use of storage capacity or managing the consumption pattern to be more friendly to the power generation, new power plants need not to be built and the efficiency of existing power plants is improved. Therefore, there is great interest to reduce the peak demand by shifting part of the peak load away from the peak time.

If large scale electricity storage were available at a feasible cost, the issue of load shifting would be resolved by implementing electrical storage. However, electricity storage is not currently viable due to the high cost and limited scales. Since buildings account for the majority of the electricity consumption and the peak load is usually caused by building cooling which is a thermal consumption, it is possible to use the building thermal storage capacity available in the buildings' thermal mass to shift the peak load to the off-peak period, usually before the peak load period, while keeping thermal comfort as the ultimate goal. This strategy is known as pre-cooling.

Recently, demand response (DR) has become a promising approach in the electricity market and the implementation of smart grid. DR is an approach to stimulate end users to change their electricity usage from regular consumption patterns, in response to the time-varying price or time of use (TOU) price of electricity [35, 36]. Thermal storage in building thermal mass has been recognized as an important passive asset to shift demand for decades, and there has been a number of simulation and experimental studies on reducing the peak demand by adjusting temperature setpoints of cooling systems[37, 38, 39].

3.1. Existing Pre-cooling Strategies

With the time of use price of electricity, it is reasonable to adjust the cooling setpoint to be on the high end when the electricity price is high, while still maintain the conform in the buildings. To make use of the thermal mass of the buildings, one can also set the cooling setpoint to be on the low end prior to the high price time period. Since the building thermal has been pre-cooled, it would reduce the cooling need during the high price period. Several pre-programmed control strategies are introduced in [40]. The baseline night-setup strategy is applied in most of the current commercial buildings, in which the temperature set-points are set at the lower bound of comfort region during entire occupied hours and cooling is shut off during the unoccupied period. In this case, the building thermal mass is not used to take advantage of the time of use price difference of electricity. In the step-up strategy, the cooling setpoints are raised to a higher value during the peak price period, reducing the demand of electricity during this on-peak period. In addition, the cooling pre-stored in building thermal mass can be discharged. The linear-up strategy is a compromise of the previous two strategies. Xu et al. (2004) also suggest to extend the pre-cooling period to unoccupied hours to store more cooling.

Optimal and advanced control techniques for DR and building energy efficiency constantly emerge, facilitating the level of DR from manual to semi-automated and fully-automated [41]. Some methods, for example artificial intelligence-based [42] and reinforcement learning [43] are model-free but usually need large amounts of data from specific buildings, meaning that even though they have been proven successful for a particular building, the performance cannot be guaranteed for other buildings. Therefore, modeling still plays an important role in building energy control. With the help of various modeling packages, accurate modeling for large scale buildings is available [44, 45, 46].

Significant peak demand reduction has been shown by previous studies of optimal demand response control that takes into account an electricity market where a time-varying rate is applied. Pre-cooling (or preheating) is the basic action to shift peak demand away from on-peak periods [47]. In the demand limiting strategy, the zone temperature trajectories are obtained by solving an optimization problem under a pre-determined target demand during on-peak hours [48, 49, 50]. However, these methods are open-loop strategies and are not able to deal with real time disturbances such as building internal loads and weather changes.

While many studies focused on reducing the energy consumption or peak demand, there has been less work on reducing the energy and demand costs of building energy systems [51]. From the above description it is evident that a desirable demand response control strategy should accomplish the following objectives simultaneously.

1. Be able to optimize the trade-off between the energy consumption and demand cost by taking advantage of the time of use price difference, while maintaining the room temperature to be within the comfort zone during the occupied period.

2. Be able to make use of the building thermal storage to store and release cooling dynamically.

3. Be able to handle real time and predicted changes of load disturbances, weather changes, and even price changes.

3.2. Feedback and Economic MPC Strategy

Model predictive control is a natural fit for the building demand reduction to achieve the above desired objectives. However, it may not be appropriate to apply the standard MPC strategy that are used for continuous process operations for the purpose of building demand reduction for the following reasons.

- The commercial building operations have daily cycles, with only about ten hours of occupied period where temperature control is of concern. within this occupied period the price of electricity can change several times, making it necessarily a dynamic operation rather than a steady state operation.

- The building thermal dynamics is usually slow, in the order of hours to reach steady state. Therefore, it is difficult to separate in time scale the economic objectives and the building thermal dynamics. In other words, the usual hierarchy of steady state real time optimization and dynamic predictive control is not suitable for demand reduction in the thermal operation of buildings.

- The ambient temperature, which is a major disturbance to the building temperature control, hardly reaches steady states. The ambient temperature cycles in the same time scale as the electricity price change.

The intertwine of the economic objectives and the building and disturbance dynamics requires the development of an economic MPC strategy. Instead of using a quadratic control criterion as in the standard MPC [52], an economic objective function is designed as follows,

$$\min J = \sum_{t=1}^{N} [Ec(t) \cdot \Delta t \cdot P(t)] + Dc \cdot \max_{t_d \leq t \leq N} \{P(t)\} \quad (2)$$

where J denotes the total electricity expense which is a combination of energy and demand costs. t_d is the time when demand charge kicks in, $Ec(t)$ accounts for the time-of-use electricity rate and Dc is the demand charge rate. For example, a rate plan offered by Southern California Edison (SCE) [53] divides a day into on-peak, mid-peak and off-peak periods. It is particularly designed for medium-sized commercial and industrial customers. $\Delta t = 0.25$hr is the time interval and N is the total number of time steps per day.

Eq.(2) is a min-max optimization problem. Ma et al. [54, 55] convert the minimax problem into a linear program so that it can be solved by a linear programming routine. Although the power consumption $P(k)$ is related to the economic objective, it does not indicate whether the comfort level of the building is achieved. To do so we must make the economic objective be subject to the dynamic thermal model constraint and the temperature of various zones or rooms of the building be subject to the comfort constraint. A dynamic power and temperature model can be built as follows.

$$\begin{aligned} P(k) &= G_P(q)T_{sp}(k) + H_P(q)d(k) + V_P(k) \\ T(k) &= G_T(q)T_{sp}(k) + H_T(q)d(k) + V_T(k) \end{aligned}$$

where $T(k)$, $T_{sp}(k)$, and $d(k)$ are measured zone temperature, zone temperature setpoints for HVAC, and measured disturbances. $V_P(k)$ and $V_T(k)$ are the effect of unmeasured disturbances on the power and zone temperature models. These models are derived from system identification in Ma et al. [54, 55], where specific care must be taken in designing the setpoint perturbations. The MPC constraints are set up as follows.

$$T_{min}(k) \leq T(k) \leq T_{max}(k) \tag{3}$$

$$T_{sp,min}(k) \leq T_{sp}(k) \leq T_{sp,max}(k) \tag{4}$$

$$P_{min}(k) \leq P(k) \leq P_{max}(k) \tag{5}$$

The constraint (5) allows one to implement the demand limiting strategy.

With the inequality constraints and the equality constraints, the optimization problem is formulated as a linear program, which can be solved by the Matlab built-in function *Linprog*. In each time step, only the current temperature setpoints $T_{sp}(k)$ in the optimal solution is implemented. This optimization procedure is repeated and a new problem will be formulated in subsequent time steps when new measurement data are available. Forecast of the measured and unmeasured disturbances with the MPC model or an external weather forecast model can be incorporated in the economic MPC problem, Making it possible to implement feedforward control actions.

3.3. A Simulation Case Study

Ma et al.[54, 55] report an economic MPC results of a single story commercial building located in Chicago, Illinois modeled in EnergyPlus. Shown in Fig. 5, the building is divided into five air-conditioned zones which include one interior and four exterior zones. A set of controllable actuators and temperature sensors is installed in these zones. The impact of building cooling loads such as occupants, lighting and electrical equipment is included in the EnergyPlus model. The Energyplus software is also capable of simulating the external

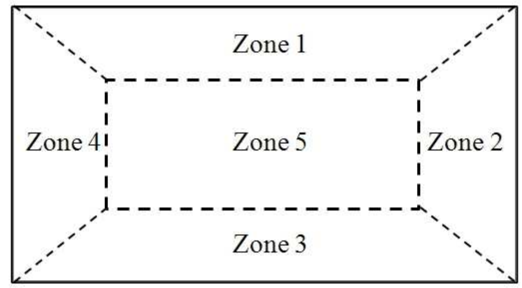

Figure 5. Five zone division floor plan

loads. A weather file that contains historical measurements of ambient temperature, relative humidity and solar radiation is incorporated.

In the MPC formulation, the zone temperature can be regulated by real-time constraints. In this work one day is divided into five periods as described in Fig. 6. The time periods are determined so that the zone temperature level in each period can be maintained within an appropriate range, rather than at a constant temperature setpoint.

1. Period 1 ($t_1 \sim t_2$): The building can be pre-cooled at as low as 18°C from the early morning until the occupied period starts. Cooling is expected to be stored in the building thermal mass and released later when necessary.

2. Period 2 ($t_2 \sim t_3$): During the off-peak and mid-peak occupied hours, zone temperature is maintained in lower half of the thermal comfort range 21 °C ∼ 23 °C with the hope that the stored cooling can be saved for utilizing in on-peak period.

3. Period 3 ($t_3 \sim t_4$): Zone temperature is free as long as within the comfort range. The stored cooling in building envelope can be either supplied or released.

4. Period 4 ($t_4 \sim t_5$): Maintain zone temperature in 23 °C ∼ 25 °C with the contribution of stored cooling.

5. Period 5 ($t_5 \sim t_1$ of the next day): Shut down the cooling system to avoid needless energy consumption.

Since t_2 and t_5 are fixed to the beginning and end of the occupied hours, there are three parameters (t_1, t_3 and t_4) that can be adjusted to set the upper and lower bounds of temperature constraints. Manipulating t_3 and t_4 can be interpreted from Fig. 6 as cutting off the areas of A and B from the thermal comfort region. Note that

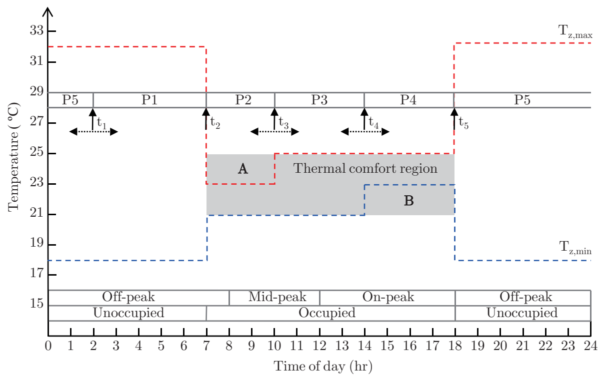

Figure 6. Five period division in MPC constraint formations.

Table 2. Ambient temperature of the simulation week (°C)

July	Sun. 1	Mon. 2	Tue. 3	Wed. 4	Thu. 5	Fri. 6	Sat. 7
Ave.	24.0	26.8	28.9	22.4	19.8	21.7	25.6
Hi.	31.7	31.1	34.4	26.1	25.0	27.8	32.2
Lo.	14.7	19.4	23.3	17.2	13.3	11.7	16.7

changing t_4 should not affect the control results much because the MPC controller tends to adjust the zone temperature trajectories to approach the upper bound of comfort region during the on-peak period.

The best scenario is selected to be $t_1 = 2$, $t_3 = 12$ and $t_4 = 12$, and this selection is not affected by the weather conditions. The ambient temperature in that simulation week varies a lot as indicated in Table 2. Corresponding simulation results of zone temperature and power profiles are shown as Fig. 7. It can be observed that the peak loads have been shifted away from the on-peak period and the on-peak power profile has been flattened. The impact of ambient temperature to energy consumption can also be seen clearly. For example, the power spike at about 3 a.m. on Tuesday was caused by the relatively warm night that made the building thermal mass more difficult to be pre-cooled. Less power was consumed over weekend than weekdays due to the different schedules of internal loads (occupants).

The performance of MPC in saving electrical costs is compared with the baseline night-setup strategy (BL), the linear-up (LU) and the step-up (SU) [49], in which setpoints are set to the lower bound of comfort region until on-peak hour begins, and then raised with a linear and step pattern respectively. Savings in energy and costs are shown in Table 3. It can be seen that MPC brings significant savings comparing to the pre-programmed control strategies.

Table 3. Weekly savings compared to the baseline

Strategy	Energy saving (%)	Cost saving (%)
Linear-up	15.29	17.42
Step-up	21.49	24.35
MPC	25.31	28.52

4. Future Work

The results described in Section 2 this paper clearly show that increased energy efficiency in the area of building control is achievable by using advanced control approaches. The existing results also point to possible improvements in the directions of improved control and fault-detection and isolation and fault-tolerant control of cooling units. In particular, recent advances in the area of identifying dynamic data-based models that utilize the strength of rigorous statistical techniques such as principal component analysis and partial least squares in conjunction with the idea of utilizing multiple linear models to capture process nonlinearity [56] can be utilized in the context of modeling of cooling units for the purpose of control. Furthermore, the area of fault-detection and isolation and fault-tolerant control has been severely understudied in the context of building systems. Results on fault-detection and isolation and reconfiguration-based fault-tolerant control [57] as well as the recently developed safe-parking approaches [58] could be adapted to significantly benefit the area of building control and lead to significant savings and improved control.

While demand reduction has been formulated as an economic MPC problem and successfully converted into a linear program, several issues remain to be tested for future work. The most important issue is to demonstrate in a real building or a building complex that the pre-cooling strategy can bring in significant cost saving and/or demand reduction. The challenge includes the ability to build a suitable dynamic model from system identification. Since detailed simulation models are sometimes available, model reduction is an alternative approach to deriving the MPC model. Another direction of future work is to incorporate other source of sensor measurements, including occupant load measurement by motion sensors and video signals. The economic

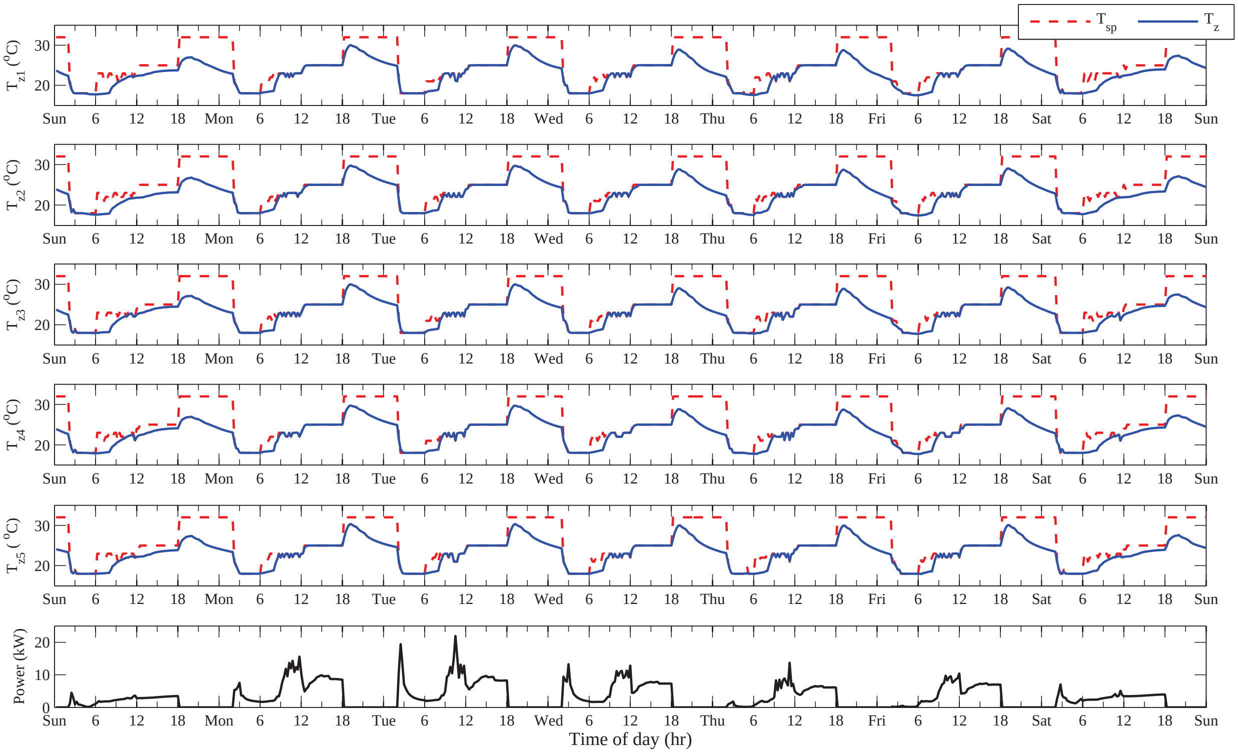

Figure 7. Zone temperature and power profiles in the weekly simulation

MPC problem could also incorporate additional objectives such as air quality and humidity control. With suitable measurements and models for these quality variables, additional savings in energy cost are expected by tight control of these quality variables.

5. Conclusions

This paper describes the application of advanced control strategies to energy systems in buildings. Two new control methods were presented. The first focused on vapor-compression cycle systems that are used in many residential and small sized buildings and are responsible for a large portion of total energy costs. Small improvements in the efficiency of operation of these systems can therefore translate into significant cost savings. An MPC strategy was used that directly handled inherent interactions in the system enabling improved setpoint tracking and disturbance rejection. The strategy also included economic optimization whereby the compressor speed was used as a proxy for energy use. Important system constraints were also integrated in the control strategy rather than being implemented separately as is the case in most current systems. The control objective was to satisfy setpoint tracking requirements within the constraint bounded region of operation and at the same time minimize the energy used by the compressor. Results showed significant improvements over the current state of the art.

The second control strategy was designed to address the seemingly intractable problem of how to reduce energy costs when electricity charges include peak demand charges during the on-peak period. Demand reduction has been formulated as an economic model predictive control problem, which is appropriate for the nature of the unsteady disturbances and dynamic price changes that happen in the same time scale as the building thermal dynamics. In the proposed MPC algorithm, the min-max optimization problem is transformed into a linear program which is solved at each time step. The proposed method aims at minimizing the costs on the daily basis, where a shrinking horizon is chosen. This shrinking horizon also allows to eventually handle dynamic pricing cases, in which the electricity rate is released by utility at the beginning of the day based on load forecast. It was shown by simulation that under the time-of-use electrical pricing structure, MPC brings substantial cost savings for a simulated building. Challenges remain in testing the control strategy in real buildings, large scale buildings, and a group of buildings as a complex.

In summary, both of the control methods presented here show promise as a way to reduce energy costs in buildings through smarter and more integrated control. However, there are still a number of issues that need to be resolved before the technology is able to be adopted more broadly. One problem is that both methods require a model of the system that is being optimized. Obtaining a model is not only costly but it is also not clear whether an accurate enough model could be obtained under practical conditions. Both approaches fit linear models to systems that are non-linear and also time-varying. Further research needs to address the issue of model accuracy and ascertain by how much the use of inevitably inaccurate models will affect the performance of the control strategies. Robustness also needs to be analyzed because of the fact that linking several previously independent control loops makes it possible that control failure would affect much larger portions of building operation than would previously have been the case. Safeguards against these system-wide failures therefore need to be developed before deployment. Finally, the issues of tuning and set-up need to be carefully researched in the context of the low cost buildings environment, and assumption of minimal expertise available on-site.

References

[1] UnderWood, *HVAC Control Systems: Modelling, Analysis and Design*. London: E & FN Spon, 1999.

[2] N. Tudoroiu, M. Zaheeruddin, and J. A. Coll, "Fault detection and diagnosis of valve actuators in hvac systems," in *Proceedings of 2005 IEEE Conference on Control Applications*, 2005, pp. 1281–1286.

[3] Y. Shen, W. J. Cai, and S. Li, "Normalized decoupling control for high-dimensional MIMO processes for application in room temperature control HVAC systems," *Control Engineering Practice*, vol. 18, pp. 652–664, 2010.

[4] I. Yang, M. S. Yeo, and K. W. Kim, "Application of artificial neural network to predict the optimal start time for heating system in building," *Energy Conversion and Management*, vol. 44, pp. 2791–2809, 2003.

[5] T. I. Salsbury and B. Chen, "A new sequencer controller for multistage systems of known relative capacities," *HVAC&R Research*, vol. 8, pp. 403–428, 2002.

[6] J. E. Seem, "A new pattern recognition adaptive controller with application to HVAC systems," *Automatica*, vol. 34, pp. 969–982, 1998.

[7] S. Katipamula and M. R. Brambley, "M.R. methods for fault detection, diagnostics, and prognostics for building systems - a review, part I," *HVAC&R Research*, vol. 11, pp. 3–25, 2005.

[8] ——, "M.R. methods for fault detection, diagnostics, and prognostics for building systems - a review, part II," *HVAC&R Research*, vol. 11, pp. 169–187, 2005.

[9] L. Mises, *Economic Calculation In The Socialist Commonwealth*. 1990 Re-print by Ludvig von Mises Institute, AL. USA, 1920.

[10] J. Rawlings and B. Stewart, "Coordinating multiple optimization-based controllers: New opportunities and challenges," *J. Proc. Contr.*, vol. 18, pp. 839–845, 2008.

[11] J. M. Gordon and K. C. Ng, "Thermodynamic modeling of reciprocating chillers," *J. App. Phys.*, vol. 75, pp. 2769–2774, 1994.

[12] T. Q. Qureshi and S. A. Tassou, "Variable-speed capacity control in refrigeration systems," *App. Therm. Eng.*, vol. 16, no. 2, pp. 103–113, 1996.

[13] A. Pacheco-Vega, M. Sen, K. T. Yang, and R. L. McClain, "Neural network analysis of fin-tube refrigerating heat exchanger with limited experimental data," *Int. J. Heat & Mass Trans.*, vol. 44, pp. 763–770, 2001.

[14] R. N. N. Koury, L. Machado, and K. A. R. Ismail, "Numerical simulation of a variable speed refrigeration system," *Int. J. Refrig.*, vol. 24, no. 2, pp. 192–200, 2001.

[15] K. C. Ng and X. L. Wang, "Thermodynamic methods for performance analysis of chillers," *Proc. Inst. Mech. Eng. Part E-J. Proc. Mech. Eng.*, vol. 219, no. E2, pp. 109–116, 2005.

[16] W. Shi, S. Shao, X. Li, and Q. Yan, "Simulation model for complex refrigeration systems based on two-phase fluid network - Part I: Model development," *Int. J. Refrig.*, vol. 31, no. 3, pp. 490–499, 2008.

[17] ——, "Simulation model for complex refrigeration systems based on two-phase fluid network - Part II: Model application," *Int. J. Refrig.*, vol. 31, no. 3, pp. 500–509, 2008.

[18] B. P. Rasmussen and A. G. Alleyne, "Control-oriented modeling of transcritical vapor compression systems," *J. Dyn. Sys. Meas. & Contr.- Trans. ASME.*, vol. 126, no. 1, pp. 54–64, 2004.

[19] B. P. Rasmussen, A. G. Alleyne, and A. B. Musser, "Model-driven system identification of transcritical vapor compression systems," *IEEE Trans. Contr. Syst. Tech.*, vol. 13, no. 3, pp. 444–451, 2005.

[20] B. D. Eldredge, B. P. Rasmussen, and A. G. Alleyne, "Moving-Boundary Heat Exchanger Models with Variable Outlet Phase," *J. Dyn. Sys. Meas. Cont.-Trans. ASME*, vol. 130, 2008.

[21] B. Li and A. G. Alleyne, "A dynamic model of a vapor compression cycle with shut-down and start-up operations," *Int. J. Refrig.*, vol. 33, no. 3, pp. 538–552, 2010.

[22] A. Alleyne, B. R. M. Keir, and B. Eldredge, "Advances in energy systems Modeling and control," in *Proceedings of the American Control Conference, New York, NY*, 2007, pp. 3658–3668.

[23] M. S. Elliott and B. P. Rasmussen, "On reducing evaporator superheat nonlinearity with control architecture," *Int. J. Refrig.*, vol. 33, pp. 607–614, 2010.

[24] M. Wallace, R. McBride, S. Aumi, P. Mhaskar, T. Salsbury, and J. House, "Energy efficient model predictive

temperature control," *Chem. Eng. Sci., in press*.

[25] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, pp. 789 – 814, 2000.

[26] P. Mhaskar, N. H. El-Farra, and P. D. Christofides, "Predictive control of switched nonlinear systems with scheduled mode transitions," *IEEE Trans. Automat. Contr.*, vol. 50, pp. 1670–1680, 2005.

[27] ——, "Stabilization of nonlinear systems with state and control constraints using Lyapunov-based predictive control," *Syst. & Contr. Lett.*, vol. 55, pp. 650–659, 2006.

[28] P. Mhaskar, "Robust model predictive control design for fault-tolerant control of process systems," *Ind. & Eng. Chem. Res.*, vol. 45, pp. 8565–8574, 2006.

[29] M. Mahmood and P. Mhaskar, "Enhanced stability regions for model predictive control of nonlinear process systems," *AIChE J.*, vol. 54, pp. 1487–1498, 2008.

[30] M. S. Elliott and B. P. Rasmussen, "Model-based predictive control of a multi-evaporator vapor compression cooling cycle," in *Proceedings of the American Control Conference*, 2008, pp. 1463–1468.

[31] ——, "A Model-Based Predictive Supervisory Controller for Multi-Evaporator HVAC Systems," in *Proceedings of the American Control Conference*, 2009, pp. 3669–3674.

[32] D. Leducq, J. Guilpart, and G. Trystram, "Non-linear predictive control of a vapour compression cycle," *Int. J. Refrig.*, vol. 29, pp. 761–772, 2006.

[33] B. P. Rasmussen and A. G. Alleyne, "Gain Scheduled Control of an Air Conditioning System Using the Youla Parameterization," *IEEE Trans. Contr. Sys. Tech.*, vol. 18, no. 5, pp. 1216–1225, 2010.

[34] "Feedback Structures for Vapor Compression Cycle Systems," M. Keir and A. Alleyne, Eds. New York City: American Control Conference, July 2007, pp. 5052–5058.

[35] M. Albadi and E. El-Saadany, "Demand response in electricity markets: An overview," in *Power Engineering Society General Meeting, 2007. IEEE*, 2007, pp. 1 –5.

[36] F. Rahimi and A. Ipakchi, "Overview of demand response under the smart grid and market paradigms," in *Innovative Smart Grid Technologies (ISGT), 2010*, 2010, pp. 1–7.

[37] J. E. Braun, "Reducing energy costs and peak electrical demand through optimal control of building thermal mass," *ASHRAE Transactions*, pp. 264–273, 1990.

[38] A. Rabl and L. K. Norford, "Peak load reduction by preconditioning buildings at night," *International Journal of Energy Research*, vol. 15, no. 9, 1991.

[39] G. P. Henze, "Energy and cost minimal control of active and passive building thermal storage inventory," *Journal of Solar Energy Engineering*, vol. 127, no. 3, pp. 343–351, 2005.

[40] P. Xu, P. Haves, M. A. Piette, and J. Braun, "Peak demand reduction from pre-cooling with zone temperature reset in an office building," 2004, university of California eScholarship Repository [http://repositories.cdlib.org/cgi/oai2.cgi] (United States) ER.

[41] S. Kiliccote, M. A. Piette, and D. Hansen, "Advanced controls and communications for demand response and energy efficiency in commercial buildings," in *2nd Carnegie Mellon Conference in Electric Power Systems: Monitoring, Sensing, Software and Its Valuation for the Changing Electric Power Industry*, 2006.

[42] M. Krarti, "An overview of artificial intelligence-based methods for building energy systems," *Journal of Solar Energy Engineering*, vol. 125, no. 3, pp. 331–342, 2003.

[43] S. Liu and G. P. Henze, "Evaluation of reinforcement learning for optimal control of building active and passive thermal storage inventory," *Journal of Solar Energy Engineering*, vol. 129, no. 2, pp. 215–225, 2007.

[44] T. Hong and Y. Jiang, "A new multizone model for the simulation of building thermal performance," *Building and Environment*, vol. 32, no. 2, pp. 123–128, 1997, doi: DOI: 10.1016/S0360-1323(96)00045-5.

[45] D. B. Crawley, L. K. Lawrie, F. C. Winkelmann, W. F. Buhl, Y. J. Huang, C. O. Pedersen, R. K. Strand, R. J. Liesen, D. E. Fisher, M. J. Witte, and J. Glazer, "Energyplus: creating a new-generation building energy simulation program," *Energy and Buildings*, vol. 33, no. 4, pp. 319–331, 2001, doi: DOI: 10.1016/S0378-7788(00)00114-6.

[46] Z. Yu, F. Haghighat, B. C. M. Fung, and H. Yoshino, "A decision tree method for building energy demand modeling," *Energy and Buildings*, vol. 42, no. 10, pp. 1637–1646, 2010, doi: DOI: 10.1016/j.enbuild.2010.04.006.

[47] P. Xu and P. Haves, "Case study of demand shifting with thermal mass in two large commercial buildings," ser. 2006 Winter Meeting of the American Society of Heating, Refrigerating and Air-Conditioning Engineers, vol. 112 PART 1, Chicago, IL, 2006, pp. 572–580.

[48] K.-H. Lee and J. E. Braun, "A data-driven method for determining zone temperature trajectories that minimize peak electrical demand," *ASHRAE Transactions*, vol. 114, no. 2, p. 65(10), 2008.

[49] K.-h. Lee and J. E. Braun, "Development of methods for determining demand-limiting setpoint trajectories in buildings using short-term measurements," *Building and Environment*, vol. 43, no. 10, pp. 1755–1768, 2008, doi: DOI: 10.1016/j.buildenv.2007.11.004.

[50] K.-H. Lee and J. E. Braun, "Evaluation of methods for determining demand-limiting setpoint trajectories in buildings using short-term measurements," *Building and Environment*, vol. 43, no. 10, pp. 1769–1783, 2008, doi: DOI: 10.1016/j.buildenv.2007.11.003.

[51] G. P. Henze, M. J. Brandemuehl, C. Felsmann, H. Cheng, A. R. Florita, and C. E. Waters, "Optimization of building thermal mass control in the presence of energy and demand charges," *ASHRAE Transactions*, vol. 114, no. 2, p. 75(10), 2008.

[52] C. V. Rao and J. B. Rawlings, "Linear programming and model predictive control," *Journal of Process Con-*

trol, vol. 10, no. 2-3, pp. 283–289, 2000, doi: DOI: 10.1016/S0959-1524(99)00034-7.

[53] SCE, “Business rate summary,” Southern California Edison, 2008, [http://www.sce.com/NR/rdonlyres/6B523AB1-244D-4A8F-A8FE-19C5E0EFD095/0/090202-Business-Rates-Summary.pdf].

[54] J. Ma, S. J. Qin, B. Li, and T. Salsbury, “Economic model predictive control for building energy systems,” in *IEEE PES Conference on Innovative Smart Grid Technologies (ISGT)*, Anaheim, CA, 2011.

[55] J. Ma, J. Qin, T. Salsbury, and P. Xu, “Demand reduction in building energy systems based on economic model predictive control,” *Chemical Engineering Science*, pp. 1–9, 2011.

[56] S. Aumi and P. Mhaskar, “Integrating data-based modeling and nonlinear control tools for batch process control,” *AIChE J.*, 2011, (in press).

[57] P. Mhaskar, C. McFall, A. Gani, P. D. Christofides, and J. F. Davis, “Isolation and handling of actuator faults in nonlinear systems,” *Automatica*, vol. 44, pp. 53–62, 2008.

[58] R. Gandhi and P. Mhaskar, “Safe-parking of nonlinear process systems,” *Comp. & Chem. Eng.*, vol. 32, pp. 2113–2122, 2008.