Evaluation of optimal control for active and passive building thermal storage

Gregor P. Henze a,*, Clemens Felsmann b, Gottfried Knabe b

a University of Nebraska – Lincoln, Architectural Engineering, 1110 South 67th Street PKI-203D, Omaha, NE 68182-0681, USA
b Technical University of Dresden, Institute of Thermodynamics and Building Systems Engineering, Helmholtzstrasse 14, 01069 Dresden, Germany

Received 11 October 2002; accepted 11 June 2003

Abstract

Cooling of commercial buildings contributes significantly to the peak demand placed on an electrical utility grid. Time-of-use electricity rates encourage shifting of electrical loads to off-peak periods at night and weekends. Buildings can respond to these pricing signals by shifting cooling-related thermal loads either by precooling the building’s massive structure or by using active thermal energy storage systems such as ice storage. While these two thermal batteries have been engaged separately in the past, this paper investigates the merits of harnessing both storage media concurrently in the context of optimal control. The objective function is the total utility bill including the cost of heating and a time-of-use electricity rate without demand charges. The evaluation of the combined optimal control assumes perfect weather prediction and plant modeling, which justifies the application of a consecutive time block optimization that optimizes 24 hour horizons sequentially. The analysis shows that the combined utilization leads to cost savings that is significantly greater than either storage but less than the sum of the individual savings. The findings reveal that the cooling-related on-peak electrical demand of commercial buildings can be drastically reduced and justify the development of a predictive optimal controller that accounts for uncertainty in predicted variables and modeling mismatch in real time.

© 2003 Elsevier SAS. All rights reserved.

1. Introduction

The equipment and systems providing thermal comfort and indoor air quality for commercial buildings consume 42% of the total energy used in buildings [1]. Energy use and utility cost can be reduced significantly by increasing the efficiency of this equipment, by distributing thermal energy more efficiently and by more closely meeting the needs of building occupants. The energy efficiency of system components for heating, ventilating, and air-conditioning (HVAC) has improved considerably over the past 20 years. For example, shipment-weighted energy efficiency ratios of unitary air conditioners in the United States have increased by 54% [2]. The average efficiency of centrifugal chillers improved by 36% and the efficiency of the best chillers increased by 50% [3]. With similar improvements in the efficiencies of boilers, motors, fans, and pumps, outstanding opportunities exist for reducing energy use and cost in commercial sites. Yet, these opportunities depend on effective building operations: e.g., a building with coincident heating and cooling due to inferior control loop parameters wastes energy regardless of boiler and chiller efficiency.

In contrast to energy conversion equipment, less improvement has been achieved in thermal energy distribution, storage and control systems in terms of energy efficiency and peak load reduction potential. Advancements are also needed to improve thermal storage systems, improve control systems and improve systems integration from a whole building perspective while meeting occupant comfort and performance requirements [4]. This paper illustrates some of the recent advancements toward these goals. In the definition of this article, ‘active’ denotes that thermal storage systems, such as ice storage, require an additional fluid loop to charge and discharge the storage tank or to deliver cooling to the existing chilled water loop. Building thermal capacitance is ‘passive’ since it requires no additional heat exchange fluid in addition to the conditioned air stream.

This paper evaluates the merits of combined optimal control of both passive building thermal capacitance and active thermal energy storage systems to minimize an objec-
tive function of choice including total energy consumption, energy cost, occupant discomfort, or a combination of these. Instead of merely satisfying instantaneous building cooling requirements, both active and passive storage inventories can be effectively harnessed in the framework of supervisory control:

(a) To exploit the performance benefits of cooler ambient conditions during nighttime for central chilled water plants, allowing for optimal scheduling of chillers, cooling towers, fans and pumps;
(b) To shape the next day’s cooling load profile by precooling the building’s massive structure at night;
(c) To make best possible use of the cost savings potential offered by conventional time-of-use and dynamic utility rate structures, including real-time pricing options that are offered by an increasing number of utilities.

Several investigators have identified promising savings potentials when building operation has been optimized in buildings without storage [5–10]. Moreover, recent analyses suggest significant performance merits from either active [11–21] or passive [22–33] thermal storage inventory under optimal control.

The combined use of both storage media under optimal control has been investigated for a 24-hour deterministic simulation study which revealed that significant operating cost savings (∼18%) and electrical demand reduction can be achieved [34]. Optimal building control proved most effective in dry climates with large diurnal temperature swings, in the presence of utility rates strongly encouraging load-shifting, and when cool storage systems allow more effective load-shifting than building precooling alone. These results suggest the investigation of combined optimal storage utilization facilitated by a predictive supervisory controller suitable for implementation in commercial buildings. This paper lays the groundwork for such a closed-loop model-based predictive optimal controller by investigating an overall solution approach that can be employed in real time.

Two essential assumptions are applied:

(a) Weather, occupancy, non-cooling electrical loads are perfectly predicted.
(b) The building thermal response is perfectly represented by the building model, i.e., there is no mismatch between the modeled and actual building behavior.

Given these assumptions, closed-loop optimal control is not necessary here as updated forecasts do not offer superior information and a consecutive time block optimization approach (described further below) is applied instead. The evaluation of the potential utility cost savings for a wide range of parameters will be documented in a future article.

2. Description of the analysis

2.1. Investigated building

We investigate a three-story office building as shown in Fig. 1 with five thermal zones per floor, i.e., 15 thermal zones in total. The perimeter zones have an area of 288 m² each, while the core zone has an area of 576 m². Total area per floor is thus 1728 m² and the building total is 5184 m². Counting the exterior envelope, floor, and ceiling surfaces, the building mass is approximately 770 kg·m⁻² of floor area, thus can be considered heavy-weight construction.

Peak building occupancy is 10 m²·person⁻¹. Each office worker contributes 132 W of internal gain, where 54% are assumed to be sensible and 46% latent. Peak lighting density is 20 W·m⁻². The occupancy and lighting schedules for a weekday are shown in Fig. 2, where hour 13 refers to the hour from 12 to 13. On weekends and holidays building occupancy is zero and lighting density is 5% of the peak value.

The office building was first modeled in EnergyPlus [35] and subsequently a TRNSYS [36] model was derived and validated. For a series of identical days (July 21 in Phoenix, Arizona from TMY2 weather data), good agreement of the zone temperature and cooling load profiles for both dynamic simulation programs was achieved. Subsequent
annual analysis revealed a building design cooling load of 470 kW.

The building is equipped with a central chilled water plant with a capacity of CCAP_{base} = 250 kW including a thermal energy storage system with a capacity of SCAP = 2500 kWh and a second dedicated chiller with a capacity of CCAP_{TES} = 250 kW. Thus, the base chiller is downsized by 47% and the active TES tank can meet the peak load alone for 5.3 hours. The base chiller has a constant coefficient-of-performance (COP) of 4.5 and the dedicated TES chiller has constant COP of 3.0. The zones are conditioned using a variable air volume (VAV) air-handling unit with hot water reheat at the VAV terminal boxes. Outside air intake is controlled by an economizer cycle using return air temperature limit.

2.2. Base case

We will state cost savings relative to a “base case”, which is a chilled water system that experiences the same cooling load and weather profiles and uses the same HVAC systems subject to the same utility rate structure as the corresponding optimized storage system. The active TES system is governed by the chiller-priority control strategy, i.e., the base chiller is used to serve the building cooling load up to its capacity CCAP_{base}, while the active storage is used to meet the cooling loads exceeding CCAP_{base}. The passive building thermal storage inventory is not utilized; During occupancy, a cooling zone setpoint of 24°C and a heating setpoint of 20°C is maintained; during unoccupied times, the HVAC systems are turned off and the temperatures are allowed to float.

The performance metric for all cases is the total utility cost for operating the office building over a selected time horizon, which includes electricity and heating costs. The electrical utility rate structures includes time-of-use differentiated energy charges (S-kW h\(^{-1}\)), while the utility rate for purchased heating is considered constant.

2.3. Passive thermal storage system modeling

The building structure responds to changes in zone temperature setpoints \( T_{Z,SP} \). The zone temperature \( T_z \) is directly affected only by the net convective heat flux according to the discrete-time energy balance on the zone air mass

\[
C_Z \frac{\Delta T_Z}{\Delta t} = \sum_{i} Q_{conv,i}
\]

(1)

where \( C_Z \) is the zone thermal capacitance. These convective heat fluxes include contributions from interior wall surfaces due to transmission and delayed release of solar gains, HVAC systems, internal gains, as well as infiltration. Of those, the current interior wall surfaces fluxes depend on a history of past inside and outside air and surface temperatures as well as inside and outside heat fluxes. The transient response of the building envelope is typically modeled by transforming the heat diffusion equation

\[
\frac{\partial T_Z}{\partial t} = \alpha \frac{\partial^2 T_Z}{\partial x^2}
\]

(2)

(where \( \alpha \) is the thermal diffusivity) into a conduction transfer function (CTF), where the inside and outside surface heat fluxes are determined with the help of construction-specific CTF coefficients \( a, b, c, \) and \( d \).

\[
\dot{q}_{s,o} = \sum_{k=0}^{n_a} a_k T_{s,o,t-k\Delta t} - \sum_{k=1}^{n_b} b_k T_{s,i,t-k\Delta t} - \sum_{k=1}^{n_d} d_k \dot{q}_{s,o,t-k\Delta t}
\]

(3)

The zone temperature setpoints can be varied between 15 and 30°C during unoccupied periods and between 20 and 24°C during occupied periods. Building precooling reduces the convective contributions from inside surfaces during occupied periods by depressing the average envelope temperature during unoccupied periods.

2.4. Active thermal storage system modeling

The defining feature of any storage system is its ability to bridge a temporal gap between supply and demand. In an active thermal energy storage system, the temporal occurrence of electrical cooling-related loads can be separated from that of the thermal (cooling) loads. Fig. 3 shows that the building cooling load can be met by any combination of contributions from the base chiller and the active TES system, while the dedicated TES chiller only serves to charge the active TES.

Changes in the state-of-charge \( x \) of the active TES system are described in discrete time by

\[
x_{k+1} = x_k + \Delta x_k
\]

(4)

subject to the state constraints

\[
x_{\min} \leq x \leq x_{\max} = 1
\]

(5)

![Fig. 3. Central chilled water plant configuration.](image-url)
Table 1

Modes of operation of chilled water plant

<table>
<thead>
<tr>
<th>Mode</th>
<th>Mode</th>
<th>TES charge/discharge rate</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM1</td>
<td>Discharging</td>
<td>( u \leq 0 )</td>
<td>( Q_{dis} = -u \frac{CAP}{\Delta t} )</td>
</tr>
<tr>
<td>PM2</td>
<td>Charging</td>
<td>( u &gt; 0 )</td>
<td>( Q_{dis} = 0; Q_{ch} = u \frac{CAP}{\Delta t}; Q_{base} = Q_L - Q_{dis} )</td>
</tr>
</tbody>
</table>

where \( u_k \) is the dimensionless TES charge/discharge rate subject to its own nonlinear constraints

\[
\alpha_{kmin} \leq u_k \leq \alpha_{kmax}
\]

(6)

The charge and discharge capacities depend on the available thermal energy storage inventory and current cooling load. The constraints on the control variable \( u \) are formulated as

\[
\alpha_{kmin} = \max \left\{ -\frac{\dot{Q}_L,k}{\Delta t}, \lambda_{kmin} - \lambda_k \right\}
\]

(7)

and

\[
\alpha_{kmax} = \min \left\{ CCAP_{TES} \frac{\Delta t}{\Delta SCAP}, \lambda_{kmax} - \lambda_k \right\}
\]

(8)

Thus, no actions can be taken that would lead to states-of-charge outside feasible limits, i.e., full and empty storage tank, respectively. Further, no more than the current load can be discharged and the TES chiller capacity \( CCAP_{TES} \) limits the maximum charge rate \( \alpha_{kmax} \). There is no explicit ice or chilled-water tank model and heat transfer limitations on the charging and discharging rates are not considered, i.e., we assume an idealized loss-free thermal battery.

Depending on the current cooling load, a choice of active TES charging/discharging rate \( u \) determines the mode of operation of the central chilled water plant as shown in Table 1.

2.5. Optimal control modeling

2.5.1. Monthly cost function

Optimal control is defined as that control trajectory that minimizes the total monthly utility bill \( C_m \) for electricity and heating:

\[
J_m = \min C_m = \min \left\{ C_{elec,m} + C_{heat,m} \right\}, \text{ where } \\
C_{elec,m} = C_{energy,m} + C_{demand,m} = \sum_{k=1}^{K_m} r_{e,k} P_k \Delta t_k \\
+ \max_{1 \leq k \leq K_m} \left\{ r_{d,k} P_k \right\}
\]

(9)

\[
C_{heat,m} = \sum_{k=1}^{K_m} r_h \dot{Q}_{head,k} \Delta t_k
\]

where \( r_{e,k} \) and \( r_{d,k} \) are the energy and demand rates for electricity according to the utility tariff in effect for time \( k \), \( K_m \) is the number of hours in the current month, \( P_k \) is the total facility electricity demand, \( \Delta t_k \) is a time increment of one hour, \( r_s \) is the unit cost of heat delivered, and \( \dot{Q}_{head,k} \) is the heating demand in hour \( k \). For the analysis presented here, load-shifting to off-peak hours is encouraged only through a substantial energy rate differential; demand rates are not considered and the cost function simplifies to

\[
J_m = \min C_m = \min \left\{ C_{energy,m} + C_{heat,m} \right\}
\]

(10)

2.5.2. Consecutive time block optimization

Consecutive time block optimization (CTBO) is employed, i.e., the predictive optimal controller carries out an optimization over a predefined planning horizon \( L \) and the complete generated optimal strategy is executed. At any time \( k^* \), the required external variables (such as weather information) are predicted over a planning horizon \( L \) and the optimal policy that minimizes \( J_L \) is determined. The complete strategy is executed without correcting for improved forecasts available during \( k^* < k < k^* + L \). After \( L \) time step the process is repeated. The planning horizon is \( L = 24 \) hours throughout this study.

The alternative approach is closed-loop optimization (CLO), i.e., the predictive optimal controller carries out an optimization over a predefined planning horizon \( L \) and of the generated optimal strategy only the first action is executed. At the next time step the process is repeated. The final control strategy of this near-optimal controller over a total horizon of \( K \) steps is thus composed of \( K \) initial control actions of \( K \) optimal strategies of horizon \( L \), where \( L < K \). By moving the time window of \( L \) time steps forward and updating the control strategy after each time step, a new forecast is introduced at each time step and yields a policy which is different from the policy found without taking new forecasts into account.

In the limiting case of perfect forecasts, both CLO and CTBO can be expected to produce identical results. When the future is subject to uncertainty, i.e., in the case of an actual implementation, CLO-based predictive optimal control is expected to exhibit superior performance. Since the focal point of this paper is to identify the relative performance of jointly optimizing the active and passive building thermal storage, we assume perfect predictions and use CTBO.

The optimal solution \( J_L \) found at current time \( k^* \) is associated with \( L \) global temperature setpoints \( [T_{Z,SP}]_{k^*+L} \) and \( L \) active TES charge/discharge rates \( [u_k]_{k^*+L} \).

2.5.3. Sequential optimization and building modes

The cost of electrical energy \( C_{energy,L} \) is affected by both the active and passive building thermal storage strategy. The choice of zone temperature setpoints will affect the cooling load, which has to be known for the active storage to be controlled properly. Therefore, there is a causal relationship
from the passive to the active storage, which requires us to solve the passive storage first, followed by the optimization of the active thermal storage inventory on the basis of the previously determined optimal building cooling load profile.

Due to the presence of simple upper and lower zone temperature bounds, the passive thermal storage (building mass) component of the control problem proved to be solved effectively with the help of a common implementation of the quasi-Newton method, which is described below. The use of a direct search method (Nelder–Mead Simplex) led to an excessive number of function evaluations (TRNSYS runs) because of cost penalties arising from bound violations. To reduce the numerical complexity of the passive storage optimization problem, a simplification is introduced: Instead of optimizing \( L \) variables, only one global zone setpoint \( T_{Z,SP} \) is determined for each combination of occupancy (occupied, unoccupied) and utility rate periods (on-peak, off-peak), defined as building mode (BM), occurring over the next \( L \) time steps.

BM1: Unoccupied and off-peak rates;
BM2: Unoccupied and on-peak rates;
BM3: Occupied and off-peak rates;
BM4: Occupied and on-peak rates.

During each building mode, the corresponding control variable is kept constant as shown in Fig. 4(a). Since these few variables describe stepped profiles for each control variable, we denote them as solution parameters \( SP \). For the given occupancy and utility rate periods and assuming hourly time steps, the solution space for an \( L = 24 \) hour horizon is reduced from 24 dimensions to 5 dimensions. For any horizon \( L \), the number of parameters can increase or decrease depending on how many distinct occupancy and rate periods are covered. Though this simplification causes the solution to become slightly suboptimal compared to the full solution, the problem now becomes computationally tractable.

The active storage (TES) optimization problem is characterized by complex and nonlinear constraints as expressed by Eqs. (7) and (8), yet simple state transitions as characterized by Eq. (4). This class of problem is most readily solved using dynamic programming, which is described below, and yields \( L \) solution variables as shown in Fig. 4(b).
2.5.4. Iterative sequential optimization

Fig. 5 illustrates how the least utility cost \( J_L \) over horizon \( L \) is determined. At time zero and starting with initial zone temperature setpoints \([T_{z,sp}]_\text{init}\) halfway between the upper and lower bounds and no active storage utilization \([u]_\text{init} = 0\) the passive storage inventory is optimized to minimize \( C_L \). As a result, the optimal building cooling load profile is computed and handed over to the active storage optimization, which calculates an optimal TES charge/discharge strategy. In a second pass, the optimal active storage utilization strategy and the previously found optimal zone temperature setpoint profile are employed to determine the new optimal temperature setpoint profile and optimal utility cost \( J_L \). This cycle is repeated until the optimal cost \( J_L \) converges. Typically, convergence is attained after 2–3 iterations. Previously optimal solutions are stored as starting values for subsequent optimizations to reduce execution time.

2.6. Optimization algorithms

We investigate two classes of optimization algorithms: a quasi-Newton method, which approximates the function gradient through finite differences, and dynamic programming for sequential decision making problems. Among those methods that utilize gradient information, quasi-Newton methods are the most popular. They collect curvature information on the cost function at each iteration to describe a quadratic model problem

\[
\min_x \left\{ \frac{1}{2} x^T H x + c^T x + b \right\} \tag{11}
\]

where the Hessian matrix, \( H \), is a positive definite symmetric matrix, \( c \) is a constant vector, and \( b \) is a constant. The optimal solution \( x^* \) occurs when the partial derivatives of \( x \) vanish, i.e.,

\[
\nabla f(x^*) = Hx^* + c = 0 \quad \Rightarrow \quad x^* = -H^{-1}c \tag{12}
\]

Newton-type methods calculate the Hessian \( H \) directly, which is numerically very demanding. Quasi-Newton methods avoid the direct computation of the Hessian by extracting curvature information from observed behavior \( f(x) \) and \( \nabla f(x) \) and subsequently approximating the Hessian numerically [37]. We employ the popular method by Broyden, Fletcher, Goldfarb, and Shanno (BFGS):

\[
H_{i+1} = H_i + \frac{q_i q_i^T}{s_i^T s_i} - \frac{H_i s_i s_i^T H_i}{s_i^T H_i s_i} \tag{13}
\]

where \( s_i = x_{i+1} - x_i \) and \( q_i = \nabla f(x_{i+1}) - \nabla f(x_i) \). In the presented case, the gradient information is derived by partial derivatives using numerical differentiation via finite differences: Each decision variable \( x \) is perturbed and the rate of change in the cost function is determined. Then at each iteration \( i \), a line search is performed in the direction of \( d = -H_{i}^{-1} \cdot \nabla f(x_i) \) (14)

The task of minimizing operating cost using active thermal storage inventory is framed as a sequential decision-making process of decision variable \( u \). The optimization technique dynamic programming commonly used for this type of problems was first formally introduced by the mathematician Richard Bellman in 1957. Bellman’s Principle of Optimality [38] states that:

“An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.”

In other words, the optimal solution to an \( L \)-step process must come from the optimal solution of an \( L - 1 \)-step process that is based on the optimal outcome of the first step. The solution of one \( L \)-step process will thus be found recursively by optimizing \( L \) single-step processes in reverse
time by starting at the end of time and moving back to “now”. To apply, the cost function has to be incrementally additive and the dynamic system has to be discrete.

3. Results

The utility rate is assumed to be $0.20/kWh on-peak and $0.05/kWh off-peak; no demand charge is levied. The on-peak period is weekdays from 9 AM to 6 PM, off-peak all remaining hours. The building is occupied from 7 AM to 5 PM.

The viewgraphs in this section are created on the basis of simulations in which July 21 in Phoenix, AZ is repeated over and over again until steady-state conditions are attained after
Table 2

System sizing for investigated control strategies

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Optimization Units</th>
<th>Sizing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Base case WITHOUT active storage</td>
<td>CCAP_{base} kW 500, CCAP_{tes} kW 0, SCAP kWh 0</td>
</tr>
<tr>
<td></td>
<td>Base chiller fully sized, no active storage; night setup</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Base case WITH active storage</td>
<td>CCAP_{base} kW 250, CCAP_{tes} kW 250, SCAP kWh 2500</td>
</tr>
<tr>
<td></td>
<td>Base chiller downsized; chiller-priority active storage control; night setup</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Passive-only</td>
<td>CCAP_{base} kW 500, CCAP_{tes} kW 0, SCAP kWh 0</td>
</tr>
<tr>
<td></td>
<td>Base chiller fully sized, no active storage; zone setpoints optimized</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Active-only</td>
<td>CCAP_{base} kW 500, CCAP_{tes} kW 250, SCAP kWh 2500</td>
</tr>
<tr>
<td></td>
<td>Base chiller fully sized, optimal active storage control; night setup</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Active and passive</td>
<td>CCAP_{base} kW 500, CCAP_{tes} kW 250, SCAP kWh 2500</td>
</tr>
<tr>
<td></td>
<td>Base chiller fully sized, optimal active storage control; zone setpoints optimized</td>
<td></td>
</tr>
</tbody>
</table>

Case 1 represents the base case in which cooling loads have to be met without any storage available. Case 2 makes use of active thermal storage as governed by chiller-priority control, i.e., the downsized base chiller meets the cooling loads up to its capacity CCAP_{base}, thereafter the active storage contributes the remainder. The dedicated active storage chiller requires SCAP/CCAP_{tes} = 10 hours to recharge an empty storage tank. Case 3 optimizes the passive storage capacity by properly precooling the building structure using a fully sized base chiller. In case 4, the active storage is now optimized instead of governed by a simple rule such as chiller-priority. Finally, case 5 optimizes both active and passive storage media and represents the focus of this research.

Case 5 is solved by optimizing each 24 hour interval sequentially, i.e., as a series of consecutive time blocks (CTBO) of 24 hours length each. The CTBO method does not allow for the consideration of newly available new information as it becomes available. However, it represents a reference scenario for comparison as we assume perfect prediction for this study.

The thick lines in Fig. 6 represent the upper and lower temperature bounds for the operation of the office building on a weekday. It can be seen how passive-only control decides on substantial nighttime precooling down to about 21°C zone temperature averaged over all 15 zones. When the temperatures are allowed to float, the average zone temperature rises beyond 28°C during unoccupied times. The average zone temperature rises beyond 28°C during unoccupied times. The

Fig. 6. Average zone temperature profiles.
combined utilization of active and passive storage leads to less precooling than in the passive-only case. All strategies involving passive storage allow for the temperatures to float from the end of occupancy at 5 PM to 6 PM because electricity prices are still high (on-peak) during this time. After 6 PM, electricity prices are low and the building is unoccupied.

The inventory of the active storage is shown in Fig. 7 from midnight to midnight for those strategies involving active storage. For the base case with active storage under chiller-priority control, the storage is fully charged during off-peak hours and discharged by about 50% during the day. The active-only optimization discharges fast as of 8 AM, but slows down during the early afternoon hours to end up empty by the end of occupancy. The combined storage utilization approach makes less use of the active storage.

Fig. 8 illustrates the effect of precooling on the daytime cooling load profile and shows how the building cooling load is shifted away from the expensive on-peak period to the off-peak period for all cases involving passive storage utilization. The passive-only approach leads to the lowest on-peak cooling loads, next comes the CTBO approach to the combined case.

Reducing on-peak electrical demand is a side effect of shifting expensive on-peak cooling loads to off-peak periods for energy-only optimizations as can be seen in Fig. 9. While the base case with active storage under chiller-priority control already reduces the demand by 20%, the combined
Fig. 9. Total building electrical demand profiles.

Fig. 10. Total hourly building operating cost profiles.

Table 3
Summary of daily operating costs

<table>
<thead>
<tr>
<th></th>
<th>Total building operating cost</th>
<th>HVAC hourly operating cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base case without TES</td>
<td>Base case with TES</td>
</tr>
<tr>
<td></td>
<td>$347.42</td>
<td>$314.97</td>
</tr>
<tr>
<td>Savings</td>
<td>BC without TES: 16.4%</td>
<td>BC with TES: 7.8%</td>
</tr>
<tr>
<td></td>
<td>BC with TES: 16.8%</td>
<td>BC with TES: 8.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BC with TES: 26.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BC with TES: 18.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BC with TES: 19.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BC with TES: 20.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BC with TES: 57.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BC with TES: 46.5%</td>
</tr>
</tbody>
</table>
optimization cuts the overall demand nearly in half. Active-only and passive-only are both superior to the base case with active storage, but inferior to the combined case solved by CTBO.

For a utility rate without demand charges, we can plot daily profiles of utility cost. The total hourly building operating cost including non-cooling cost is shown in Fig. 10. The areas under each curve represent the total daily operating cost. It is obvious that on-peak cost savings are traded off against nighttime expenses for recharging active and/or passive storage inventories.

Fig. 11 illustrates how the cooling related costs are effectively shifted to nighttime periods. In fact, the combined storage cases lead to near-zero cooling costs during the on-peak period.

Finally, Table 3 provides an overview of the daily cost savings achieved for this prototypical day in Phoenix, AZ. Based on total utility cost, savings of about 16% can be achieved for either passive- or active-only storage, and about 26% for the combined case when compared to the base case without storage. Compared to the base case with active storage under chiller-priority control, savings of about 8% can be achieved for either passive- or active-only storage, and about 18% for the combined case. Based on cooling related utility cost only, savings of about 37% can be achieved for either passive- or active-only storage, and about 57% for the combined case when compared to the base case with active storage under chiller-priority control, savings of about 20% can be achieved for either passive or active-only storage and about 46% for the combined case. These results show that given strong load-shifting incentives, the benefits of the proposed optimization system may be substantial.

4. Conclusions and future work

This study investigated the potential of building thermal storage inventory, in particular the combined utilization of active and passive inventory, for the reduction of electrical utility cost using common time-of-use rate differentials. The findings reveal that when an optimal controller is given perfect weather forecasts and when the building model used for predictive control perfectly matches the actual building, utility cost savings and on-peak electrical demand reductions are substantial. While this work established the theoretical maximum performance, future efforts are required to determine how strongly prediction performance and model mismatch deteriorate the controller performance. Eventually, once an acceptable weather predictor is available and system identification routines calibrate the underlying model, lab and field experimentation will be required to verify these savings figures during actual operation.

Acknowledgements

The authors gratefully acknowledge the support of this work through US Department of Energy Cooperative Agreement No. DE-FC26-01NT41255.

References

