MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

A Method for Computing Optimal Set-Point Schedule for HVAC Systems

Nikovski, D.; Xu, J.; Nonaka, M.

TR2013-050 June 2013

Abstract

We propose a method for model-based control of building air conditioning systems that minimizes energy costs while maintaining occupant comfort. The method uses a building thermal model in the form of a thermal circuit identified from collected sensor data, and reduces the building thermal dynamics to a Markov decision process (MDP) whose decision variables are the sequence of temperature set-points over a suitable horizon, for example one day. The main advantage of the resulting MDP model is that it is completely discrete, which allows for a very fast computation of the optimal sequence of temperature set-points. Experiments on thermal models demonstrate savings that can exceed 50% with respect to usual control strategies in buildings such as night setup.

2013 REHVA World Congress (CLIMA)

A Method for Computing Optimal Set-Point Schedules for HVAC Systems

Daniel Nikovski[#1], Jingyang Xu[#2], and Mio Nonaka[*3]
[#]*Mitsubishi Electric Research Laboratories,*
201 Broadway, Cambridge, MA 02139, USA
{[1]nikovski,[2]jxu}@merl.com
[*]*Mitsubishi Electric Corporation,*
8-1-1, Tsukaguchi-Honmachi, Hyogo 661-8661, Japan
[3]nonaka.mio@dc.mitsubishielectric.co.jp

Abstract
We propose a method for model-based control of building air conditioning systems that minimizes energy costs while maintaining occupant comfort. The method uses a building thermal model in the form of a thermal circuit identified from collected sensor data, and reduces the building thermal dynamics to a Markov decision process (MDP) whose decision variables are the sequence of temperature set-points over a suitable horizon, for example one day. The main advantage of the resulting MDP model is that it is completely discrete, which allows for a very fast computation of the optimal sequence of temperature set-points. Experiments on thermal models demonstrate savings that can exceed 50% with respect to usual control strategies in buildings such as night setup.

Keywords - load shifting; optimal control; embedded Markov chains; dynamic programming; Markov decision processes

1. Introduction

Heating, ventilation and air conditioning (HVAC) systems are responsible for the largest part of electrical energy consumption in most developed countries, and optimization of their energy use is one of the most promising means for reducing greenhouse gases and the overall cost and pollution associated with generation of electrical energy. A simple strategy for reducing energy consumption by HVAC systems is to increase the temperature set-point to the highest value that is still comfortable for building occupants (e.g., 28^oC). While this strategy would be more cost-effective in comparison to cases when the zone set-point is held constant at a lower value, it is not necessarily the most efficient strategy. Another possible strategy is to vary the set-point throughout the day so as to shift the load from peak periods when the HVAC system is less efficient and energy costs might be higher (e.g., the afternoon in summer) to off-peak periods when the system is more efficient and energy is less expensive (e.g., in the morning). This strategy effectively results in pre-cooling (respectively, pre-heating in winter) of the building, and is based on the ability of the thermal mass of the building to

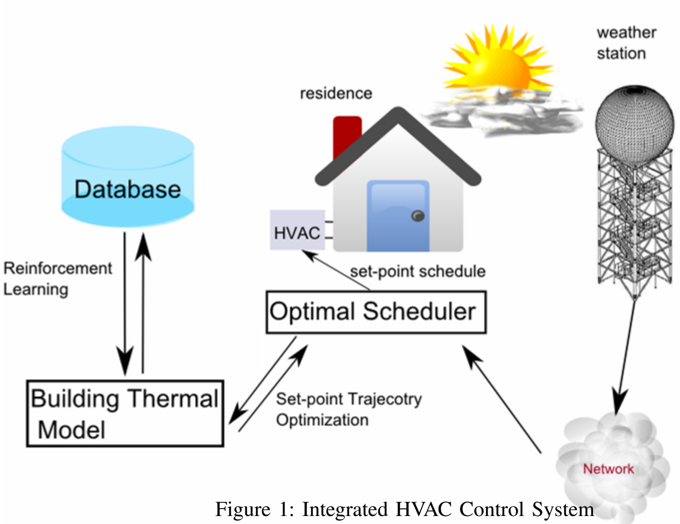

Figure 1: Integrated HVAC Control System

store heat over time intervals with duration of several hours, comparable to the time difference between peak and off-peak periods. The question then emerges of how to compute the optimal sequence of set-points throughout an entire day. This is a sequential optimization problem that could be solved if a suitable building model and a computational procedure that can use this model can be found. We propose such a procedure based on Markov decision processes (MDP) that uses a building thermal model in the form of a third-order thermal circuit. Section 2 describes the formulation of the sequential optimization problem, as well as the thermal circuit, and Section 3 explains how this problem is reduced to an MDP that can be solved efficiently. Section 4 presents experimental results in comparison with two widely used practical scheduling strategies for energy saving, and Section 5 concludes and suggests directions for further improvements to the model and computational procedure.

2. Set-Point Optimization for HVAC Systems

2.1. Overall Operation of the Optimizing System

The general operation of the optimizing method and system is shown in Fig. 1. The system constantly collects operational and environmental temperature data from the sensors on the indoor and outdoor HVAC units, stores the measurements in a database, and uses a learning algorithm to acquire a predictive building thermal model that can predict future temperatures of the controlled zone given its current state, outdoor conditions, and a proposed sequence of temperature set-points that implicitly control the amount of heat transferred by the HVAC system. A typical prediction horizon is 24 hours, with prediction/control step a fraction of an hour, for example 5 minutes. This means that the predictive model is iterated multiple times to produce predictions for the entire prediction horizon. Since future outdoor temperatures that will determine the heat load on the HVAC system are not known in advance, weather forecasts from public meteorological services (e.g. the National Weather Service of the USA) can be retrieved over the Internet and used for prediction. The job of the optimal set-point scheduler is to find the optimal sequence of set-points for the entire optimization period. However, typically, the entire sequence is not executed; rather, after the optimal sequence is found, only its first value is assigned to the thermostat set-point, and at the next time steps, the operation is repeated, resulting in continuous rolling-horizon operation. If computational resources are limited, the optimal sequence can also be re-computed less frequently, for example once every hour.

The computational complexity of the optimization problem of finding the optimal set-point sequence could be very large. For a 24-hour optimization horizon and 5-minute prediction/control step, the sequence would have $K = 288$ different set-points. If we choose a discrete set of A possible values for the set-points at any given time, there is an exponential number of possible set-point sequences — A^K — and enumeration and evaluation of all such sequences is not computationally feasible. This paper describes a computationally efficient procedure for finding the optimal sequence without exhaustive enumeration, based on the modeling formalism of Markov decision process (MDP) models and dynamic programming. The building thermal model, the HVAC efficiency function, and the MDP model used in the computational procedure are described below.

2.2. Building Thermal Model

The method uses a building thermal model in the form of a third-order thermal circuit ([2], [4]), which is a popular modeling formalism that strikes a reasonable balance between physical realism of the model, prediction accuracy, and model order. The state vector x of the model is defined as the three temperatures $x \doteq [T_Z, T_{Isurf}, T_{Osurf}]^T$,where T_Z is the zone temperature (under the assumption of well mixed air with uniform temperature), T_{Isurf} is the temperature of the inside surface of the wall surrounding the zone, and T_{Osurf} is the temperature of the outside wall that is in direct contact with the outside air that has temperature $T_{Outside}$, assumed to be known at all times. Of the three components of the state vector, only the first one, T_Z, is assumed to be measurable; the other two are hidden (unobservable) state variables. The dynamic behavior of the thermal circuit can be expressed by means of the following set of ordinary differential equations (ODE), [1],

[2]:

$$\begin{aligned}
\frac{dT_{Osurf}}{dt} &= \frac{T_{Ooutdise}-T_{Osurf}}{R_{Eo}C_{Eo}} + \frac{T_{Isurf}-T_{Osurf}}{R_{Em}C_{Eo}} + \frac{S_1}{C_{Eo}} \\
\frac{dT_{Isurf}}{dt} &= \frac{T_{Osurf}-T_{Isurf}}{R_{Em}C_{Ei}} + \frac{T_Z-T_{Isurf}}{R_{Ei}C_{Ei}} + \frac{S_2}{C_{Ei}} \\
\frac{dT_Z}{dt} &= \frac{T_{Outside}-T_Z}{R_{Win}C_Z} + \frac{T_{Isurf}-T_Z}{R_{Ei}C_Z} + \frac{S_3}{C_Z}
\end{aligned} \tag{1}$$

In this ODE set, S_1, S_2, and S_3 are heat gains, and dynamics are parametrized by means of thermal resistances R and capacitances C. However, in most HVAC systems, the amount of heat transferred by the system is not directly controllable; instead, the usual means of controlling such systems is to specify a temperature set point by means of a thermostat. Still, for any existing controller, we can describe the evolution of the system by the general set of equations $x_{k+1} = f_k(x_k, a_k)$, where x_k is the state of the system at time t_k, $a_k = T_{S,k}$ is the temperature set-point specified at time t_k, f_k is a state evolution function (not necessarily linear), and the system evolves in discrete time such that $t_k = k\Delta t$. In general, the function f_k is time dependent, because it is parametrized by the uncontrollable inputs $T_{Outside}$, O, and V, all of which vary throughout the day. The goal of the optimization algorithm is to select a sequence of temperature set-points $T_{S,1}$, $T_{S,2}$, $\ldots$, $T_{S,K}$ such that a cumulative performance measure (energy or cost) dependent on the states traversed by the system and the controls applied to it is optimized over the entire optimization horizon of K steps:

$$J = \sum_{k=1}^{K-1} g_k(x_k, a_k) + h(x_K) \tag{2}$$

where g_k is the running cost for a single period (again, time dependent), and h is a terminal cost associated with the final state x_K. Regardless of the feedback controller used to regulate zone temperature starting from a given system state for a specified target set-point, the amount of heat Q transferred by the air conditioner can be computed for any period. From there, the amount of electrical power W needed to transfer this amount of heat can obtained by means of the coefficient of performance (COP) of the system, defined as $COP \doteq |Q|/W$. As discussed previously, the efficiency (COP) of an HVAC system based on a vapor compression cycle depends strongly on the outdoor temperature. Furthermore, if the objective is to minimize total monetary cost, rather than energy, that cost can also be readily estimated by multiplying the amount of energy by its price obtained from time-of-use tariffs. Following this procedure the running cost functions g_k can be estimated.

It is much more difficult to estimate the terminal cost h at the end of the optimization period. One possible simplification would be to assume that $h(x_K) = 0$ for all states x_K. This simplification would bias the scheduler towards set-point sequences that leave the building as hot as possible, but such solutions are not unreasonable for office spaces that are occupied only until the end of the official business hours — intuitively, if the indoor air temperature reaches the upper limit of the comfort zone in the very last period, more energy would be saved than if the temperature at that time was lower. For this reason, the choice $h(x_K) = 0$ should be acceptable in practice.

3. Markov Decision Process Model

Methods for minimizing the cumulative cost 2 for arbitrary functions f, g, and h do not generally exist, although solutions for special cases, such as linear dynamics and quadratic costs, have been known and used for a long time [7]. However, in this case, even though the thermal circuit model could be represented by a linear system, the function f_k is neither linear nor time independent, due to the time-varying driving conditions $T_{Outside}$, O, and V, and the operation of the HVAC controller. Moreover, the cost functions g_k and h are generally not quadratic in the state and control, since the COP curve can have an arbitrary form for different HVAC devices. In such cases, the optimal control has to be found by means of either numerical methods ([5]) or reinforcement learning ([6]). An alternative strategy for solving the optimal control problem is to convert the continuous-state-space dynamical system into a Markov decision process (MDP) with discrete state space and solve it by means of existing algorithms for MDP. The proposed algorithm is one such method.

A discrete-space MDP is described by the tuple (S, A, p, r), where S is a discrete set of states such that the MDP occupies one of the states $s_k \in S$ at any time t_k. A is a discrete set of actions $a \in A$, and the transition probability function $p(s_{k+1}|s_k, a_k)$ expresses the probability of being in state s_{k+1} at time t_{k+1} if the MDP was in state s_k at time t_k and control (action) a_k was applied at that time. $r(s_k, a_k)$ is a cost function that expresses the cost of applying control a_k in state s_k. The MDP evolves in discrete time, where time steps usually have the same duration ($t_k = k\Delta t$). The goal is to optimize a performance measure $R = \sum_{k=0}^{K} r(s_k, a_k)$, much like in the continuous case. We construct all transition functions of the MDP as sets of coefficients for suitably defined convex combinations, equivalent to the barycentric coordinates of the end points of transitions of the continuous dynamical system. The procedure for constructing the MDP has the following steps (explained in more detail in [8]):

1) Define the feasible subspace of the entire state space R^3 of the thermal

circuit as a cube limited by the minimal and maximal values for the three state variables T_Z, T_{Isurf}, and T_{Osurf}. A possible minimal value is the lesser of the lowest comfortable zone air temperature (e.g., 21^0C) and the lowest outdoor temperature $T_{Outside}$ throughout the day; similarly, a possible maximal value is the highest outdoor temperature $T_{Outside}$ during the optimization period. Tighter bounds for some of the state variables, for example T_Z, are also possible.

2) Sample a set of N anchor points $X = \{x^{(1)}, x^{(2)}, \ldots, x^{(N)}\}$ from the feasible subspace of the system (in this case, it is three-dimensional). Either random sampling or a regular grid can be used.
3) Triangulate the set of anchor points, for example by means of Delaunay triangulation . The result will be a set of three-dimensional simplices (tetrahedra).
4) Define a set $S = \{s^{(1,1)}, \ldots, s^{(1,N)}, s^{(2,1)}, \ldots, s^{(2,N)}, \ldots, s^{(K,N)}\}$ of $K \times N$ MDP states, and associate all K MDP states $s^{(k,i)}$, $1 \leq k \leq K$ with anchor point $x^{(i)}$. The states can be organized in K subsets of N slices each, such that the subset $S_k = \{s^{(k,1)}, s^{(k,2)}, \ldots, s^{(k,N)}\}$ defines the set of states that can be occupied by the MDP at time t_k.
5) Define a set of actions A by discretizing the temperature interval that corresponds to the comfort zone at suitable steps, for example $A = \{21^oC, 22^oC, \ldots, 28^oC\}$.
6) For every time step t_k, $1 \leq k < K$, every anchor point $x^{(i)}$, $1 \leq i \leq N$, and every action $a^{(l)}$, $1 \leq l \leq |A|$:

    a) Find the resulting building state $y = f_k(x^{(i)}, a^{(l)})$ if the HVAC system is operated with temperature set-point $a^{(l)}$ from time t_k to time t_{k+1}, starting in state $x^{(i)}$ and using the thermal model and HVAC control policy.
    b) In general, the resulting building state y will not coincide with any of the anchor states $x^{(i)}$, $1 \leq i \leq N$. However, it will always be within exactly one simplex (tetrahedron) of the Delaunay triangulation. Let $x^{(i_1)}$,$x^{(i_2)}$,$x^{(i_3)}$, and $x^{(i_4)}$ be the four anchor points corresponding to the vertices of this tetrahedron. Then, there exist four uniquely defined real numbers c_1, c_2, c_3, and c_4, such that $y = \sum_{m=1}^{4} c_m x^{(i_m)}$, $0 \leq c_m \leq 1$, and $\sum_{m=1}^{4} c_m = 1$, also known as the barycentric coordinates of y with respect to the tetrahedron.
    c) For a starting state $s_k = s^{(k,i)}$ and chosen action $a_k = a^{(l)}$, define the transition function $p(s^{(k+1,j)}|s^{(k,i)}, a^{(l)}) = Pr(s_{k+1} = s^{(k+1,j)}|s_k = s^{(k,i)}, a_k = a^{(l)})$ to the N possible successor states $s_{k+1} = s^{(k+1,j)}$, $1 \leq i \leq N$, as follows. If anchor point $x^{(j)}$ associated with MDP

state $s^{(k+1,j)}$ was one of the four vertices x_m of the tetrahedron containing $y = f_k(x^{(i)}, a^{(l)})$, then $p(s^{(k+1,j)}|s^{(k,i)}, a^{(l)}) \doteq c_m$; otherwise $p(s^{(k+1,j)}|s^{(k,i)}, a^{(l)}) \doteq 0$. That is, in effect, we assume that when executing a transition from a given MDP state, the transition can be only to one of the four states in the next time slice whose associated anchor points define the tetrahedron containing the end state y of the actual system evolution. Transitions to all other states in other times slices have zero probability, resulting in MDP transitions from one subset (slice) of MDP states to the next.

d) The transition cost $r(s_k, a_k)$ is equal to the cost experienced by the continuous system: $r(s^{(k,i)}, a^{(l)}) = g_k(x^{(i)}, a^{(l)})$.

Since the resulting MDP model has no loops (transitions are always from one time slice to the next), the value function (cost-to-go) $V(s)$ of every state s of the MDP (and from there the optimal policy $a_k = \pi^*(s_k)$) can be computed by means of dynamic programming starting from the last time slice and proceeding backward in time. Making use of the auxiliary function $q(s,a)$, the computation uses Bellman back-ups of the form:

$$q(s^{(k,i)}, a) = r(s^{(k,i)}, a) + \sum_{j=1}^{N} p(s^{(k+1,j)}|s^{(k,i)}, a)V(s^{(k+1,j)}) \tag{3}$$

$$V(s^{(k,i)}) = \min_{a \in A}[q(s^{(k,i)}, a)] \tag{4}$$

$$\pi^*(s^{(k,i)}) = argmin_{a \in A}[q(s^{(k,i)}, a)], \tag{5}$$

for all time periods $0 \le k < K$, and initializing the value function of the terminal states $s^{(K,i)}$, $1 \le i \le N$, as $V(s^{(K,i)}) = h(x^{(i)})$, where the terminal costs $h(x^{(i)})$ could be assumed to be zero for this optimization problem, as discussed above. Notably, since a Bellman back-up would be performed exactly once for each MDP state, the computational complexity of the entire solution procedure will be $O(KN)$.

However, the identified optimal policy $a = \pi^*(s^{(k,i)})$ is a mapping from the state of the MDP to the optimal action, whereas we need a mapping $a = \mu_k^*(x)$ from the state of the *building* x at time t_k instead. As noted, the state of the building x will generally not coincide with any of the N anchor points $\{x^{(1)}, x^{(2)}, \ldots, x^{(N)}\}$ at any given time, but will always be within a tetrahedron defined by four such points $x^{(i_m)}$, $1 \le m \le 4$. Then, if c_m are the corresponding barycentric coordinates of x with respect to the four anchor points, we can estimate the cost-to-go $\hat{q}_k(x,a)$ of the state/action pair (x,a)

as

$$\hat{q}_k(x,a) \doteq \sum_{m=1}^{4} c_m q(s^{(k,i_m)},a),$$

and the optimal policy would be $\mu_k^*(x) = argmin_{a\in A}[\hat{q}_k(x,a)]$. Note that the optimal policy will be time-dependent, in general.

4. Experimental Analysis

4.1. Experimental Conditions

A set of experiments in simulation were performed in cooling mode, using actual measured data for weather conditions, human activities data, etc. A single-zone building thermal model was estimated from measured data, with parameters shown in Table 1. The comfort zone for building temperature was assumed to be the interval $[21^oC, 26^oC]$. The set of possible temperature set-points for the zone thermostat was chosen to be the set $\{21^o, 21.25^o, 21.5^o, \ldots, 26^oC\}$. For the efficiency model of the air conditioner, we chose a COP curve for refrigerant R22. For the Time-Of-Use (TOU) electricity price structure, we used the same tariff as in [3], by dividing each day into a peak period and an off-peak period, with ratio of peak to off-peak prices equal to 3. Four summer weather profiles collected in Cambridge, Massachusetts in July and August of 2003 were used as weather conditions in the experiment.

For comparison, we used two known energy-saving strategies: the Night Set-up Strategy (NSS), where a set-point is specified for the official business hours, and the zone temperature is left to float freely at other times, and the Demand Limiting Strategy (DLS) that uses a fixed set-point schedule that starts pre-cooling the building 3 hours before the start of office hours, setting the temeperature set-point at the lowest comfortable value, and in the afternoon, when peak period starts, gradually raises the set point to the upper limit of the comfort zone.

4.2. Experimental Results

The results over four days of experiments are shown in Table 2. The cost for the NSS, DLS, and MDP-based set-point scheduling methods is computed according to a time-of-use tariff. It can be seen that the MDP-based scheduler can save costs significantly, sometimes exceeding 50%. Example simulation and optimization results over one day are shown in Fig. 2.

Table 1. Thermal circuit parameters

Parameter Name	Value	Description
R_{Win}	0.1295	Resistance of windows
R_{EO}	0.3846	Resistance of outside wall surface
R_{Em}	0.0511	Resistance of wall
R_{EI}	0.0261	Resistance of inside wall surface
C_{EO}	7.3447e+05	Capacitance of outside wall surface
C_{EI}	9.5709e+05	Capacitance of wall
C_Z	9.03473e+04	Capacitance of inside wall surface

Table 2. Simulation results over four days, starting from the same initial condition $x_0 = [24.2, 24.3, 24.9]^T$, but with four different outdoor temperature profiles. Energy (E) is in kWh. The monetary cost (C) is computed according to a time-of-use tariff, in equivalent monetary units for the three set-point scheduling strategies, with 360 units/kWh in off-peak hours and 1080 units/kWh in peak hours.

Date	**E(MDP)**	**E(NSS)**	**E(DLS)**	**C(MDP)**	**C(NSS)**	**C(DLS)**
07 – 22	655	828	2200	948	1750	2240
07 – 28	405	642	1800	476	1450	1850
08 – 16	1030	1210	2380	1710	2630	2450
08 – 28	1290	1600	2410	2390	3760	2480

5. Conclusion

A general method for controlling building zone air temperature by setting temperature set-points was proposed in this paper. The method uses a low-order thermal circuit model and converts the continuous-state building dynamics into a discrete Markov decision problem, relying on the similarities between the mathematical properties of convex combinations (barycentric coordinates) and probability transition functions. Finding the optimal control policy for the converted MDP representation is computationally very efficient, and reduces to backward dynamic programming. This procedure is linear in the number of states of the MDP, and the overall computational time can be varied according to needs and the available computational resources by means of adjusting the number of states. The favorable computational complexity of the algorithm can be employed in a receding horizon controller for continuous re-optimization of the set-points of HVAC systems.

In a set of experiments with a single zone in cooling mode, the MDP-based scheduler proposed in this paper significantly outperforms traditional scheduling strategies. Further research will extend this method to multi-zone buildings, and will address the influence of model inaccuracy and uncertainty in input data (outdoor temperature, etc.). Another possibility for improvement of the method is to also include humidity both in the building thermal model and the definition of the comfort zone for building occupants.

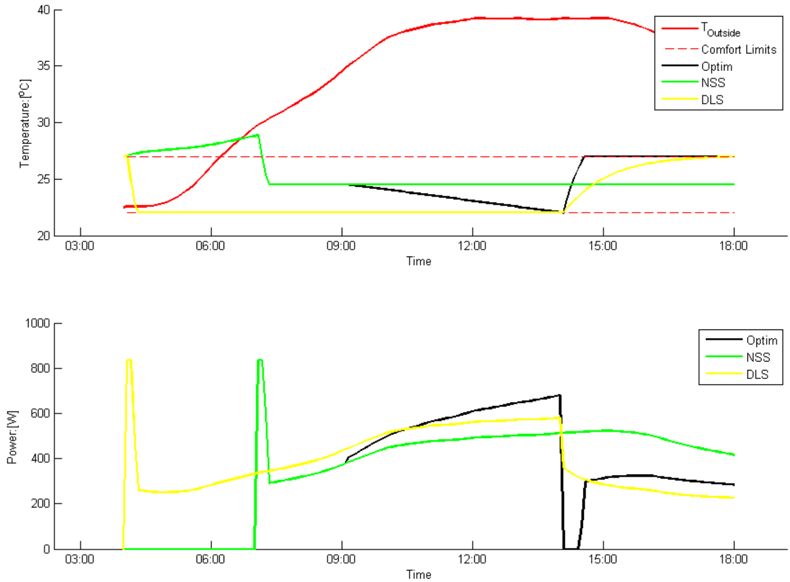

Figure 2: Simulation and optimization results over one particularly hot day. By pre-cooling before the onset of the peak pricing period (2pm), the optimized set-point sequence expends more power than the NSS strategy, but avoids operation of the HVAC systems when it is most expensive and least efficient. As a result, its total cost is 2471, vs. 3220 for the baseline NSS strategy. The DLS, although also pre-cooling, is not tailored to this particular building, and results in sub-optimal cost, too: 2726.

References

[1] M. J. Jimenez and H. Madsen, "Models for describing the thermal characteristics of building components," *Building and Environment*, vol. 43, no. 2, pp. 152–162, 2008.

[2] M. J. Jimenez, H. Madsen, and K. K. Andersen, "Identification of the main thermal characteristics of building components using matlab," *Building and Environment*, vol. 43, no. 2, pp. 170–180, 2008.

[3] J. E. Braun, "A near-optimal control strategy for cool storage systems with dynamic electric rates," *HVAC&R Research*, vol. 13, no. 4, pp. 557–580, 2007.

[4] Q. Zhou, S. Wang, X. Xu, and F. Xiao, "A grey-box model of next-day building thermal load prediction for energy-efficient control," *International Journal of Energy Research*, no. 15, pp. 1418–1431, 2008.

[5] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, Massachusetts: Athena Scientific, 2000. Volumes 1 and 2.

[6] R. S. Sutton and A. G. Barto, *Reinforcement learning*. Cambridge, Massachusetts: The MIT Press, 1998.

[7] R. F. Stengel, *Optimal Control and Estimation*. Mineola, NY: Dover, 1986.

[8] D. Nikovski and A. Esenther, "Construction of embedded Markov decision processes for optimal control of non-linear systems with continuous state spaces," in *IEEE Conference on Decision and Control and European Control Conference*, pp. 7944–7949, 2011.