CHAPTER 42

Supervisory Control Strategies and Optimization

COMPUTERIZED building and energy management and control systems provide a variety of effective ways to reduce utility costs and energy consumption associated with maintaining environmental conditions and thermal comfort in commercial buildings. These systems can incorporate advanced control strategies that respond to inputs including changing weather, building conditions, occupancy levels and utility rates to minimize operating costs, energy consumption and greenhouse gas emissions.[1]

HVAC and other building energy systems are typically controlled using hierarchical control structure where two or more levels of control, from local through to supervisory level, are combined to form a more sophisticated control system designed to achieve particular high level functions or objectives- such as maintaining temperature within a space. With this philosophy of control, controller intelligence increases from lower to higher levels within the hierarchy. The lowest control level typically exists only to provide **local-loop control** of a single set-point through manipulation of an actuator. For example, the supply air temperature discharged from a cooling coil is controlled by adjusting the opening of a valve that provides chilled water to the coil. The upper control level, typically termed **supervisory control**, specifies the set-points and other modes of operation that are time dependent and may also provide system performance monitoring capabilities. Ideally, the supervisory control level would determine optimal set-points and operation modes that result in minimized operating cost and/or energy consumption, maximizing comfort, and may also identify potential faults or alarms within the control system. Distributed control structures have also been applied to building energy systems, however further research is still required to determine the efficacy of such a structure and the benefits that this approach may provide over more traditional and well understood control systems.

The performance of large, commercial HVAC systems can be improved through better local-loop and supervisory control. Proper tuning of local-loop controllers can enhance comfort, reduce energy use, and increase component life. Systems that are properly commissioned or tuned, such as through a re-commissioning process or the use of automated fault detection and diagnostics software, will ensure that the theoretical performance gains from supervisory control strategies are realized. Set points and operating modes for cooling or heating plant equipment can be adjusted by supervisory control strategies or static optimization to maximize overall operating efficiency. Dynamic optimization strategies for ice or chilled-water storage systems can significantly reduce on-peak electrical energy and demand costs to minimize total utility costs. Similarly, thermal storage inherent in a building's structure can be dynamically controlled to minimize utility costs. In general, strategies which take advantage of thermal storage work best when dynamic optimization is applied utilizing forecasts of future energy requirements.

Significant increases in computational power and communications capabilities mean that both supervisory and distributed control systems are now able to incorporate many new data streams and simultaneously co-optimize a number of performance metrics. Rather than simply regulating temperature, the supervisory control system may manage thermal comfort while minimizing utility cost, energy consumption and greenhouse gas emissions. Ubiquitous consumer electronics such as smart-phones, tablets and laptops mean that direct feedback of occupant preferences may be obtained rather than relying on inferred statistical models. Building thermal response and

---

[1]The preparation of this chapter is assigned to TC 7.5, Smart Building Systems.

internal gain forecast models (learned by the controller) allow optimal start-up, night-purge and economy modes while utility interaction allows participation in demand response programs and cost reductions in demand and capacity charges.

Where resources are constrained– by equipment sizing, maintenance or imposed through energy targets or demand charges– a coordinated approach to resource allocation is required to ensure an equitable balance of comfort for all building occupants. This is increasingly likely to an important control scheme design consideration weather conditions change and as uptake of intermittent renewable generation requires energy users to respond to this resource variability. New energy pricing models will substantially reward users who have this flexibility.

This chapter focuses on the opportunities and control strategies associated with using supervisory control strategies and optimization methods applied to cooling systems, heating systems, air handling units and zone equipment. The chapter is divided into five major sections: (1) a methods section focused on the basic principles of control systems, supervisory control strategies, static and dynamic optimization methods; (2) applications of supervisory control and optimization to cooling systems, including cooling tower fan control, chilled water reset strategies, sequencing chillers, static optimization of cooling plant equipment, and dynamic optimization of cooling systems with discrete thermal storage, passive thermal storage, and thermally activated building systems; (3) applications to heating systems and boilers; (4) applications to air handling units, including supply air temperature and static pressure reset strategies; and (5) applications to zone controls, such as night setback.

TERMINOLOGY

Air distribution system: includes terminal units [variable-air-volume (VAV) boxes, etc.], air-handling units (AHUs), ducts, and controls. In each AHU, ventilation air is mixed with return air from the zones and fed to the cooling/heating coil.

Air-side economizer control: used to select between minimum and maximum ventilation air, depending on the condition of the outdoor air relative to the conditions of the return air. Under certain outdoor air conditions, AHU dampers may be modulated to provide a mixed air condition that can satisfy the cooling load without the need for mechanical cooling.

Building thermal mass storage: storing energy in the form of sensible heat in building materials, interior equipment, and furnishings.

Capacity: Heating or Cooling output at design or rating condition or, in certain contexts, at current operating condition.

CAV systems: air-handling systems that have fixed-speed fans and provide no feedback control of airflow to the zones. Zone temperature is controlled to a set point using a feedback controller that regulates the amount of local reheat applied to the air entering each zone.

Charging: storing cooling capacity by removing heat from a cool storage device; or storing heating capacity by adding heat to a heat storage device.

Chilled-/hot-water/steam loop: consists of pumps, pipes, valves, and controls. Two different types of pumping systems are considered in this chapter: primary and primary/secondary. With a **primary pumping** system, a single piping loop is used and water that flows through the chiller or boiler also flows through the cooling or heating coils. When steam is used, a steam piping loop sends steam to a hot-water converter, returning hot condensate back to the boiler. Another piping loop then carries hot water through the heating coils. Often, fixed-speed pumps are used with their control dedicated to chiller or boiler control. **Dedicated control** means that each pump is cycled on and off with the chiller or boiler that it serves. Systems with fixed-speed pumps and two-way cooling or heating coil valves often incorporate a **water bypass valve** to maintain relatively constant flow rates and reduce system pressure drop and pumping costs at low loads. The valve is typically controlled to maintain a fixed pressure difference between the main supply and return lines. This set point is termed the **chilled-** or **hot-water loop differential pressure**. Sometimes, primary systems use one or more variable-speed pumps to further reduce pumping costs at low loads. In this case, water bypass is not used and pumps are controlled directly to maintain a water loop differential pressure set point.

Chiller or boiler plant: one or more chillers or boilers, typically arranged in parallel with dedicated pumps, provide the primary source of cooling or heating for the system. Individual feedback controllers adjust the capacity of each chiller or boiler to maintain a specified supply water temperature (or steam header pressure for steam boilers). Additional control variables include the number of chillers or boilers operating and the relative loading for each. For a given total cooling or heating requirement, individual chiller or boiler loads can be controlled by using different water supply set points for constant individual flow or by adjusting individual flows for identical set points.

Chiller priority: control strategy for partial storage systems that uses the chiller to directly meet as much of the load as possible, normally by operating at full capacity most of the time. Thermal storage is used to supplement chiller operation only when the load exceeds the chiller capacity.

Condenser water loop: consists of cooling towers, pumps, piping, and controls. Cooling towers reject heat to the environment through heat transfer and possibly evaporation (for wet towers) to the ambient air. Typically, large towers incorporate multiple cells sharing a common sump, with individual fans having two or more speed settings. Often, a feedback controller adjusts tower fan speeds to maintain a temperature set point for water leaving the cooling tower, termed the **condenser water supply temperature**. Typically, condenser water pumps are dedicated to individual chillers (i.e., each pump is cycled on and off with the chiller it serves).

COP: Coefficient Of Performance is the heating or cooling output divided by electrical power input; may include CW pump and tower fan as well as compressor power (see also **System COP**).

Demand limiting: a partial storage operating strategy that limits the capacity of the cooling system during the on-peak period. The cooling system capacity may be limited based on its cooling capacity, its electric demand, or the facility demand.

Discharge capacity: maximum rate at which cooling can be supplied from a cool storage device.

Discharging: using stored cooling capacity by adding thermal energy to a cool storage device or removing thermal energy from a heat storage device.

Ice storage: types of ice storage systems include the following. An **ice harvester** is a machine that cyclically forms a layer of ice on a smooth cooling surface, utilizing the refrigerant inside the heat exchanger, then delivers it to a storage container by heating the surface of the cooling plate, normally by reversing the refrigeration process and delivering hot gases inside the heat exchanger. In **ice-on-coil-**

external melt, tubes or pipes (coil) are immersed in water and ice is formed on the outside of the tubes or pipes by circulating colder secondary medium or refrigerant inside the tubing or pipes, and is melted externally by circulation the unfrozen water outside the tubes or pipes to the load. **Ice-on-coil-internal melt** is similar, except the ice is melted internally by circulating the same secondary coolant or refrigerant to the load.

Load leveling: a partial storage sizing strategy that minimizes storage equipment size and storage capacity. The system operates with the refrigeration equipment running at full capacity for 24 h to meet the normal cooling minimum load profile and, when the load is less than the chiller output, the excess cooling is stored. When the load exceeds the chiller capacity, the additional cooling requirement is obtained from the thermal storage system.

Load profile: compilation of instantaneous thermal loads over a period of time, normally 24 hours.

Nominal chiller capacity: (1) chiller capacity at standard ARI (*Standard* 550/590) rating conditions, or (2) chiller capacity at a given operating condition selected for the purpose of quick chiller sizing selections.

Partial storage: a cool storage sizing strategy in which only a portion of the on-peak cooling load is met from thermal storage, with the remainder of the load being met by operating the chilling equipment.

Precooling: a thermal energy storage (TES) strategy that allows a properly designed chiller system to operate more efficiently with lower condensing temperatures (low night ambient temperature) higher evaporating temperatures (60 to 69°F [15 to 20°C] chilled water) and at or near its most efficient part-load point. Precooling can be applied to the conditioned space, or directly to mass by passing chilled air or water through building elements such as concrete floor decks.

Primary/secondary chilled- or hot-water systems: systems designed specifically for variable-speed pumping. In the primary loop, fixed-speed pumps provide a relatively constant flow of water to the chillers or boilers. This design ensures good chiller or boiler performance and, for chilled-water systems, reduces the risk of freezing on evaporator tubes. The secondary loop incorporates one or more variable-speed pumps that are controlled to maintain a water loop differential pressure set point. The primary and secondary loops may be separated by a heat exchanger. However, it is more common to use direct coupling with a common pipe.

Storage capacity: the maximum amount of cooling (or heating) that can be achieved by the stored medium in the thermal storage device. **Nominal storage capacity** is a theoretical capacity of the thermal storage device. In many cases, this may be greater than the usable storage capacity. This measure should not be used to compare usable capacities of alternative storage systems.

Storage cycle: a period (usually one day) in which a substantial charge and discharge of a thermal storage device has occurred, beginning and ending at the same state or same time of day.

Storage efficiency: the ratio of useful heating or cooling extracted during the discharge cycle to that imparted to storage during the charging cycle. One may also define an exergic efficiency that accounts for mixing and thermal destratification as well as conduction losses.

Storage inventory: the amount of usable heating or cooling capacity remaining in a thermal storage device.

Storage priority: a control strategy that uses stored cooling to meet as much of the load as possible. Chillers are operated only if the load exceeds the storage system's available cooling capacity.

Supply air temperature: temperature of air leaving an AHU or package unit. The temperature supplied to the zones may differ from the supply air temperature due to heat transfer in the ductwork and local reheat. A local-loop controller adjusts the flow of water through the air handler cooling/heating coil using a two- or three-way valve to maintain a specified set-point temperature for the air downstream of the cooling/heating coil. The value of the supply air temperature set-point will affect energy consumption and can be adjusted by the supervisory control system.

System COP: ratio of cooling rate ($kW_{thermal}$) to system input power ($kW_{electric}$) required to operate all distribution fans and pumps as well as compressors and condenser or cooling tower fans and pumps.

Temperature set point: zone temperature set points are typically fixed values within the comfort zone during the occupied time, and zone humidity is allowed to "float" within a range dictated by the system design and choice of the supply air set-point temperature. **Night setup** is often used in summer to raise the zone temperature set points during unoccupied times and reduce the cooling requirements. Similarly, **night setback** is often used in winter to lower the zone temperature set points during unoccupied times and reduce the heating requirements. Setup and setbacks can also be used during occupied periods to temporarily curtail energy consumption.

Thermal energy storage (TES): thermal energy storage systems are of three general types. **Discrete TES** uses a tank of water (usually stratified), packed bed of rock or phase-change material, or a tank of ice and water with immersed coil. **Intrinsic (or passive) TES** uses the thermal capacitance of the building fabric, such as concrete columns, beams and decks, or wall board, possibly augmented by phase-change material. **Thermally activated building systems (TABS)** cool floors or other elements directly with embedded pipes or ducts, giving storage capacity and efficiency substantially greater than can achieved by passive TES.

System capacity: maximum amount of cooling that can be supplied by the entire cooling system, which may include the chillers and the thermal storage. **Usable storage capacity** is the total amount of beneficial cooling able to be discharged from a thermal storage device. (This may be less than the nominal storage capacity, because the distribution header piping may not allow discharging the entire cooling capacity of the thermal storage device.)

Total cooling load: The integrated thermal load that must be met by the cooling plant over a given period of time.

VAV systems have a feedback controller that regulates the airflow to each zone to maintain zone temperature set point. The zone airflows are regulated using dampers located in VAV boxes in each zone. VAV systems also incorporate feedback control of the primary airflow through various means of fan capacity control. Typically, inputs to a fan outlet damper, inlet vanes, blade pitch, or variable speed motor are adjusted to maintain a **duct static pressure** set point in the supply duct, as described in Chapter 47.

METHODS

CONTROL VARIABLES

Systems and Controls

Figures 1 and 2 show schematics of typical centralized cooling and heating systems for which control strategies are presented in this chapter. For cooling systems, the strategies generally assume that the equipment is electrically driven and that heat is rejected to the environment by cooling towers. For heating systems, boilers may be fired by a variety of fuels or powered by electricity, but are typically fired from either natural gas or #2 or #6 fuel oil. However, some strategies apply to any type of system (e.g., return from night setup or setback). For describing different systems and controls, it is useful to divide the system into the subsystems depicted in Figures 1 and 2: air distribution system, chilled-/hot-water loop, chiller/boiler plant, condenser water loop. A variant of the chilled-water system shown in Figure 1 uses chilled water directly for sensible cooling by radiant panels, chilled beams, or radiant floors. Similar strategies can be applied to variable-speed controls for air-cooled chiller condenser fans.
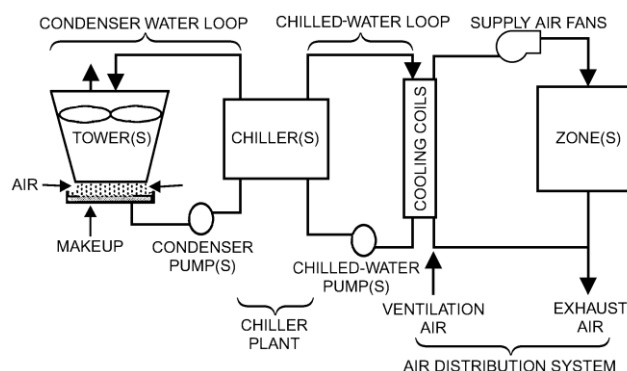

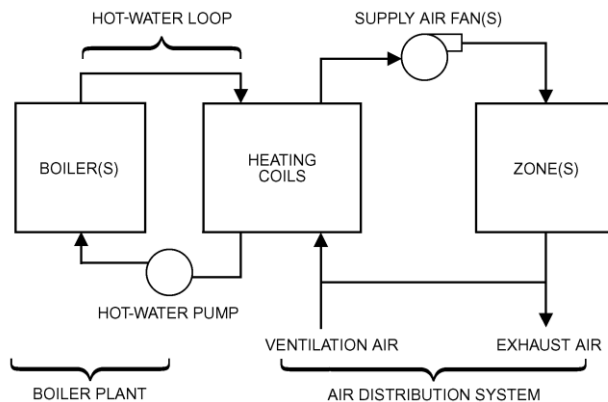**Fig. 1   Schematic of Chilled-Water Cooling System**

Fig. 2   Schematic of Hot-Water Heating System

VAV cooling system controls (Figure 1) respond to changes in building cooling requirements. As the cooling load increases, the zone temperature rises as energy gains to the zone air increase. The zone controller responds to higher temperatures by increasing local flow of cool air by opening a damper. Opening a damper reduces static pressure in the primary supply duct, which causes the fan controller to create additional airflow. With greater airflow, the supply air temperature of the cooling coils increases, which causes the air handler feedback controller to increase the water flow by opening the cooling coil valves. This increases the chilled-water flow and/or the chilled water return temperature (i.e., increases the cooling load).

Control of a radiant cooling system (a variant of Figure 1) responds directly to zone load by increasing the chilled-water flow rate. The chilled-water temperature may be controlled by the neediest zone or by an open-loop control.

Control of a hot-water heating system (Figure 2) is similar. As the heating load increases, the zone temperature falls as energy gains to the zone air decrease. The zone controller responds to lower temperatures by opening a control valve and increasing the flow of hot water through the local reheat coil. The supply airflow rate is usually maintained at its minimum value when a VAV system is in heating mode. Increasing water flow through the reheat coils reduces the temperature of the water returned to the boiler. With lower return water temperature, the supply water temperature drops, which causes the feedback controller to increase the boiler firing rate to maintain the desired supply water temperature.

For both heating and cooling, an increase in the building load results in an increase in water flow rate, which is ultimately propagated through the central system. For fixed-speed chilled- or hot-water pumps, the differential pressure controller closes the chilled- or hot-water bypass valve and keeps the overall flow relatively constant. For variable-speed pumping, the differential pressure controller increases pump speed. In a chilled-water system, the return water temperature and/or flow rate to the chillers increases, leading to an increase in the chilled-water supply temperature. The chiller controller responds by increasing the chiller cooling capacity to maintain the chilled-water supply set point (and match the cooling coil loads). The increased energy removed by the chiller increases the heat rejected to the condenser water loop, which increases the temperature of water leaving the condenser. The in-creased water temperature entering the cooling tower increases the water temperature leaving the tower. The tower controller responds to the higher condenser water supply temperature and increases the tower airflow. At some load, the current set of operating chillers is not sufficient to meet the load (i.e., maintain the chilled-water supply set points) and an additional chiller is brought online. For a hot-water system, the return water temperature and/or flow rate to the boilers decreases, leading to a decrease in the hot-water supply temperature. The boiler controller responds by increasing the boiler heating capacity to maintain the hot-water supply set point (and match the heating coil loads).

For all-electric cooling without thermal storage, minimizing power at each point in time is equivalent to minimizing energy costs. Therefore, supervisory control variables should be chosen to maximize the **coefficient of performance (COP)** of the system at all times while meeting the building load requirements. The COP is defined as the ratio of total cooling load to total system power consumption. In addition to the control variables, the COP depends primarily on the cooling load and the ambient wet- and dry-bulb temperatures. Often, the cooling load is expressed in a dimensionless form as a part-load ratio (PLR), which is the cooling load under a given condition divided by the design cooling capacity.

For cooling or heating systems with thermal storage, performance depends on the time history of charging and discharging. In this case, controls should minimize operating costs integrated over the billing period or storage cycle. In addition, safety features that minimize the risk of prematurely depleting storage capacity may be important.

SUPERVISORY CONTROL STRATEGIES

For any of the scenarios listed above, several local-loop controllers respond to load change to maintain specified set points. A supervisory controller establishes modes of operation and chooses (or resets) values of set points. At any given time, cooling or heating needs can be met with various combinations of modes of operation and set points. This chapter discusses several methods for determining supervisory control variables that provide good overall performance. These methods include resetting setpoints such as chilled water supply temperature, hot water supply temperature, supply air temperatures, differential pressure on water loops, and supply air static pressure. Another form of supervisory control includes sequencing equipment such as chillers, boilers, pumps and fans to achieve the an energy efficient mode of operation. Optimization methods may be applied to achieve the *optimal* mode of operation subject to cost, energy or comfort constraints.

Sampling Intervals for Reset Controls

Proper sampling intervals are required when resetting the set point for proportional-integral (PI) feedback control loops to prevent oscillation of the process variable in those loops. In general, the sampling time interval between reset commands should be greater than the settling time for the loop. For example, resetting both the chilled-water supply temperature set point and cooling coil discharge air temperature set point is necessary for optimal control. In this case, resetting the set points for chilled-water supply temperature and cooling coil discharge air temperature should not occur simultaneously, but at staggered intervals; the interval of reset for either loop should be greater than the settling time for the coil (the time for the discharge air temperature of the coil to reach a new steady-state temperature).

For cascaded loops, the sampling interval between reset commands should be greater than the settling time for the slower loop. An example of a cascaded loop is a VAV box controller with its flow set point determined from the space temperature of its associated zone, and the box damper controlled to maintain the flow set point. For example, in resetting static pressure set point on a variable-speed air-handler fan based on VAV box damper position, the interval between resets should be greater than the settling time of the flow control loop in a pressure-independent VAV box. (Settling time in this example is the time for the flow rate to reach a new steady-state value.)

STATIC OPTIMIZATION

Optimal supervisory control of building systems involves determining the control that optimizes selected metrics such as minimizing the total operating cost, minimizing energy consumption, or maximizing thermal comfort. Static optimization addresses the problem of optimizing the operation of a system at a given instant in time by operating each component of the system at the conditions which achieve an optimal result, such as minimal cost. Static optimization techniques applied to a general simulation can be used to determine optimal supervisory control variables. The simulation may be based on physical (Hiller and Glicksman 1977; Stoecker 1980) or empirical models. However, for control variable optimization, empirical and semiempirical models (where the model parameters are estimated from measurements) are often used. This section presents methods for static optimization useful in determining optimal control for various applications.

General Static Optimization Problem

Figure 3 depicts the general nature of the static optimization problem for a system of interconnected components. Each component in a system is represented as a separate set of mathematical relationships organized into a computer model. Its output variables and operating cost are functions of parameters and uncontrolled and controlled variables. The structure of the system of equations to be solved is dictated by the manner in which the components are interconnected.
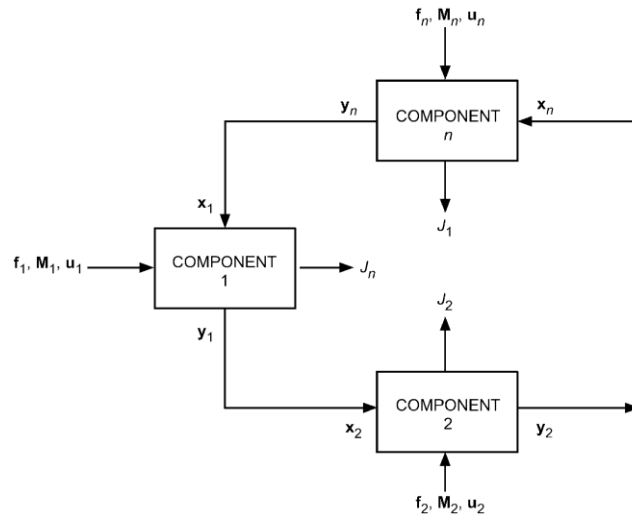
Fig. 3   Schematic of Modular Optimization Problem

A typical optimization problem is formally stated as the minimization of the sum of the operating costs of each component J_i with respect to all discrete and continuous controls, or

Minimize

$$J(\mathbf{f}, \mathbf{M}, \mathbf{u}) = \sum_{i=1}^{n} J_i \ (\mathbf{x}_i, \mathbf{y}_i, \mathbf{f}_i, \mathbf{M}_i, \mathbf{u}_i) \qquad (1)$$

with respect to $\mathbf{M}$ and $\mathbf{u}$, subject to equality constraints of the form

$$\mathbf{g}(\mathbf{f}, \mathbf{M}, \mathbf{u}) = \begin{bmatrix} \mathbf{g}_1(\mathbf{f}_1, \mathbf{M}_1, \mathbf{u}_1, \mathbf{x}_1, \mathbf{y}_1) \\ \mathbf{g}_2(\mathbf{f}_2, \mathbf{M}_2, \mathbf{u}_2, \mathbf{x}_2, \mathbf{y}_2) \\ . \\ \mathbf{g}_n(\mathbf{f}_n, \mathbf{M}_n, \mathbf{u}_n, \mathbf{x}_n, \mathbf{y}_n) \end{bmatrix} = 0 \qquad (2)$$

and inequality constraints of the form

$$\mathbf{h}(\mathbf{f}, \mathbf{M}, \mathbf{u}) = \begin{bmatrix} \mathbf{h}_1(\mathbf{f}_1, \mathbf{M}_1, \mathbf{u}_1, \mathbf{x}_1, \mathbf{y}_1) \\ \mathbf{h}_2(\mathbf{f}_2, \mathbf{M}_2, \mathbf{u}_2, \mathbf{x}_2, \mathbf{y}_2) \\ . \\ \mathbf{h}_n(\mathbf{f}_n, \mathbf{M}_n, \mathbf{u}_n, \mathbf{x}_n, \mathbf{y}_n) \end{bmatrix} \geq 0 \qquad (3)$$

where, for any component i,

$\mathbf{x}_i$ = vector of input stream variables
$\mathbf{y}_i$ = vector of output stream variables
$\mathbf{f}_i$ = vector of uncontrolled variables
$\mathbf{M}_i$ = vector of discrete control variables
$\mathbf{u}_i$ = vector of continuous control variables
J_i = operating cost
$\mathbf{g}_i$ = vector of equality constraints
$\mathbf{h}_i$ = vector of inequality constraints

Typical input and output stream variables for thermal systems are temperature and mass flow rate. Uncontrolled variables are measurable quantities that may not be controlled, but that affect component performance and/or costs, such as ambient dry- and wet-bulb temperature.

Both equality and inequality constraints arise in the optimization of chilled-water systems. One example of an equality constraint that arises when two or more chillers are in operation is that the sum of their capacities must equal the total load. The simplest type of inequality constraint is a bound on a control variable. For example, lower and upper limits are necessary for the chilled-water set temperature, to avoid freezing in the evaporator and to provide adequate dehumidification for the zones. Any equality constraint may be rewritten in the form of Equation (2) such that when it is satisfied, the constraint equation is equal to zero. Similarly, inequality constraints may be expressed as Equation (3), so that the constraint is satisfied when the LHS is greater than or equal to zero.

Braun (1988) and Braun et al. (1989a) presented a component-based nonlinear optimization and simulation tool and used it to investigate optimal performance. Each component is represented as a separate subroutine with its own parameters, controls, inputs, and outputs. The optimization problem is solved efficiently by using second-order representations of costs. Application of the component-based optimization led to many guidelines for control and a simplified system-based optimization methodology. In particular, the results showed that optimal set points could be correlated as linear functions of load and ambient wet-bulb temperature.

Zakula et al. (2011, 2012) show that, for chillers with high part-load efficiency over a wide range of compressor speed, optimal condenser fan speed and chilled-water specific flow rate are highly nonlinear functions of cooling load, and without optimal control, transport power can be much greater than compressor power at part-load and low-lift conditions. Gayeski et al. (2010a) demonstrated an empirical method for deriving near-optimal control laws from measured performance.

Cumali (1988, 1994) presented a method for real-time global optimization of HVAC systems, including the central plant and associated piping and duct networks. The method uses a building load model for the zones, based on a coupled weighting-factor method similar to that used in DOE-2 (Winkelman et al. 1993). Variable time steps are used to predict loads over a 5 to 15 min period. The models are based on thermodynamic, heat transfer, and fluid mechanics fundamentals and are calibrated to match actual performance, using data obtained from the building and plant. Pipe and duct networks are represented as incidence and circuit matrices, and both dynamic and static losses are included. Fluid energy transfers are coupled with zone loads using custom weighting factors calibrated for each zone. The resulting equations are grouped to represent feasible equipment allocations for each range of building loads, and solved using a nonlinear solver. The objective function is the cost of delivering or removing energy to meet the loads, and is constrained by the comfort criteria for each zone. The objective function is minimized using the reduced gradient method, subject to constraints on comfort and equipment operation. Optimization starts with the feasible points as determined by a nonlinear equation solver for each combination of equipment allocation. The values of set points that minimize the objective function are determined; the allocation with the least cost is the desired operation mode. Results obtained from this approach have been applied to high-rise office buildings in San Francisco with central plants, VAV, dual-duct, and induction systems. Electrical demand reductions of 8 to 12% and energy savings of 18 to 23% were achieved.

DYNAMIC OPTIMIZATION

Whereas static optimization addresses the problem of optimizing the operation of a building system at a given point in time accounting for the operating characteristics of each component, dynamic optimization addresses the problem of optimizing the control of buildings systems over time. As such, dynamic optimization must account for the possibility that future conditions, such as future weather or utility prices, may affect the optimal control decisions now and similarly control decisions now will affect operating conditions and optimal control decisions in the future. Established methods of optimal control, e.g. model-predictive control and dynamic programming, may be used to implement dynamic optimization of a building system. Specific methods of dynamic optimization will be presented relevant to the applications presented in this chapter, with a focus on cooling systems largely due to the highly weather and load dependent nature of optimal cooling system controls. Daily, monthly and annual cooling costs are particularly sensitive to time of use.

Dynamic Optimization Methods for Cooling Systems with Discrete Storage

Optimal supervisory control for cooling systems with discrete storage, such as water- or ice-storage, is a function of several factors including utility rate structure, load profile, chiller characteristics, storage characteristics, and weather. For a utility rate structure that includes both time-of-use energy and demand charges, the optimal strategy can depend on variables that extend over a monthly time scale. The overall problem of minimizing the utility cost over a billing period (e.g., a month) can be mathematically described as follows:

Minimize

$$J = \sum_{k=1}^{N}(E_k P_k \Delta\tau) + \text{Max}_{1 \le k \le N}(D_k P_k) \tag{12}$$

with respect to the control variables $(u_1, u_2, \ldots, u_N)$ and subject to the following constraints for each stage k:

$$u_{min,k} \le u_k \le u_{max,k} \tag{13}$$

$$x_k = f(x_{k-1}, u_k, k) \tag{14}$$

$$x_{min} \le x_k \le x_{max} \tag{15}$$

$$x_N = x_0 \tag{16}$$

where J is the utility cost associated with the billing period (e.g., a month); Δt is the stage time interval (typically equal to the time window over which demand charges are levied, e.g., 0.25 h); N is the number of time stages in a billing period, and for each stage k, P is the average building electrical power (kW); E is the energy cost rate or cost per unit of electrical energy ($/kWh); D is the demand charge rate or cost per peak power rate over the billing period ($/kW); u is the control variable that regulates the rate of energy removal from or addition to storage over the stage; u_{max} is the maximum value for u; u_{min} is the minimum value for u; x is the state of storage at the end of the stage; x_{max} is the maximum admissible state of storage; x_{min} is the minimum admissible state of storage; and f is a state equation that relates the state of storage at stage k to the previous state and current control.

The first and second terms in Equation (12) are the total cost of energy use and building demand for the billing period. Both the energy and demand cost rates can vary with time, but typically have two values associated with on- and off-peak periods. An even more complex

cost optimization results if the utility includes ratchet clauses in which the demand charge is the maximum of the monthly peak demand cost and some fraction of the previous monthly peak demand cost during the cooling season. With real-time pricing, the demand charge, which is the second term in Equation (12), might not exist and the hourly energy rates would vary over time according to generation and transmission costs.

For ice or chilled-water storage systems, the control variable could be the rate at which energy is added or removed from storage. In this case, the constraint given by Equation (13) arises from capacity limits of the chiller and storage heat exchanger and can also depend on the state of storage. For use of building thermal mass, the control variable could be the zone temperature(s) and the constraint of Equation (13) would be associated with comfort considerations or capacity constraints. Different comfort limits would probably apply for occupied and unoccupied periods.

In this general formulation the state equation, Equation (14), is treated as an equality constraint. The state of storage at any stage k is a function of the previous state (x_{k-1}), the control (u_k), and other time-dependent factors (e.g., ambient temperature). For lumped storage systems (e.g., ice), the state of storage can be characterized with a single-state variable. However, for a distributed storage (e.g., a building structure), multiple-state equations may be necessary to properly characterize the dynamics. The state of storage may be constrained to always lie between states corresponding to full discharge and full charge [Equation (15)]. The constraint of Equation (16) forces a steady-periodic solution to the problem. This constraint becomes less important as the length of analysis (*control horizon*) increases.

To determine a control strategy for charging and discharging storage that minimizes utility cost, Equation (12) must be minimized over the entire billing period because of the influence of the demand charge. Alternatively, the optimization problem can be posed as a series of shorter-term (e.g., daily or weekly) optimizations with a constraint on the peak demand charge according to

Minimize

$$J = \frac{\sum_{k=1}^{N} \{ E_k P_k \Delta\tau \}}{} + \text{TDC} \qquad (17)$$

with respect to the control variables (u_1, u_2,..., u_N) and a billing period target demand cost (TDC) and subject to the constraints of Equations (13) through (16) and the following additional constraint:

$$D_k P_k \leq \text{TDC} \qquad (18)$$

which arises from the form of the cost function chosen for Equation (17). At each stage, the demand cost must be less than or equal to the peak demand cost for the billing period. The peak or target demand cost TDC is an optimization variable that affects both energy and demand costs. Using Equation (17) rather than Equation (12) simplifies the numerical solution.

Two types of solutions to the optimization problem are of interest: (1) minimum billing-period operating cost and (2) minimum energy cost for a specified TDC and short-term horizon (e.g., a day). The first problem is useful for benchmarking the best control and minimum cost through simulation, but is not useful for online control because forecasts beyond a day are unreliable. Mathematical models of the building, equipment, and storage can be used to estimate load requirements, power, and state of storage. The second solution can be used for online control in conjunction with a system model and a forecaster.

For minimum operating costs (first optimization problem), $N + 1$ variables must be determined to minimize the cost function of Equation (17) over the length of the billing period. For a given value of TDC, minimization of Equation (17) with respect to the *N* charging (and discharging) control variables may be accomplished using dynamic programming (Bellman 1957) or some other direct search method. The primary advantages of dynamic programming are that it handles constraints on both state and control variables in a straightforward manner and also guarantees a global minimum. However, the computation becomes excessive if more than one state variable is needed to characterize storage. The *N*-variable optimization problem is resolved at each iteration of an outer loop optimization for TDC. Brent's algorithm (1973) is a robust method for solving the one-dimensional optimization for the demand target because it does not require derivative information. This is important because TDC appears as an inequality constraint in the dynamic programming solution and may not always be triggered.

For shorter-term optimizations (second optimization problem), dynamic programming can still be used to minimize Equation (17) with respect to the *N* charging (and discharging) control variables for a specified TDC. However, an optimal value for TDC cannot be determined when demand charges are imposed. For ice storage, Drees and Braun (1996) found that a simple and near-optimal approach is to set TDC to zero at the beginning of each billing period. Therefore, the optimizer minimizes the demand cost for the first optimization period (e.g., a day) and then uses this demand as the target for the billing period unless it is exceeded. For online optimization, the optimization problem can be resolved at regular intervals (e.g., 1 h) during each day's operation.

Ice Storage Control Optimization. Several researchers have studied optimal supervisory control of ice storage systems. Braun (1992) solved daily optimization problems for two limiting cases: minimum energy (i.e., no demand charge) and minimum demand (no energy charge). Results of the optimizations for different days and utility rates were compared with simple chiller-priority and load-limiting control strategies (see the section on Supervisory Control Strategies and Tools). For the ice-on-pipe system considered, load-limiting control was found to be near optimal for both energy and demand costs with on-peak to off-peak energy cost ratios greater than about 1.4.

Drees and Braun (1996) solved both daily and monthly optimization problems for a range of systems with internal-melt area-constrained ice storage tanks. The optimization results were used to develop rules that became part of a rule-based, near-optimal controller presented in the section on Supervisory Control Strategies and Tools. For a range of partial-storage systems, load profiles, and utility rate structures, the monthly electrical costs for the rule-based control strategy were, on average, within about 3% of the optimal costs.

Henze et al. (1997a) developed a simulation environment that determines the optimal control strategy to minimize operating cost, including energy and demand charges, over the billing period. A modular cooling plant model was used that includes three compressor types (screw, reciprocating, and centrifugal), three ice storage media (internal melt, external melt, and ice harvester), a water-cooled condenser, central air handler, and all required fans and pumps. The simulation tool was used to compare the performance of chiller-priority, constant-proportion, storage-priority, and optimal control.

Henze et al. (1997b) presented a predictive optimal controller for use with **real-time pricing (RTP)** structures. For the RTP structure considered, the demand term of Equation (17) disappears and the optimization problem only involves a 24 h period. The controller calculates the optimal control trajectory at each time step (e.g., 30 min), executes the first step of that trajectory, and then repeats that process at the next time step. The controller requires a model of the plant and storage, along with a forecast of the future cooling loads.

To apply the optimization approach described in the previous section, models for storage, system power consumption, and building loads are needed. For online optimization, simple empirical models that can be trained using system measurements are appropriate. However, physically based models would be best for simulation studies.

The optimization studies that have been performed for ice storage assumed that the state of storage could be represented with a single-state variable. Assuming negligible heat gains from the environment, the relative state of charge (i.e., fraction of the maximum available storage capacity) for any stage k is

$$x_k = x_{k-1} + \frac{u_k \Delta t}{C_s} \qquad (19)$$

where C_s is the maximum change in internal energy of the storage tank that can occur during a discharge cycle and u_k is the storage charging rate. The state of charge defined in this manner must be between zero and one.

The charging rate for storage depends on the storage heat exchanger area, secondary fluid flow rate and inlet temperature, and the thickness of ice. At any stage, the maximum charging rate can be expressed as

$$u_{k,max} = \varepsilon_{c,k,max} \dot{m}_{f,max} c_f (t_s - t_{f,i}) \qquad (20)$$

where $\varepsilon_{c,k,max}$ is the heat transfer effectiveness for charging at the current state of storage if the secondary fluid flow rate were at its maximum value of $m_{f,max}$, c_f is the secondary fluid specific heat, $t_{f,i}$ is the temperature of secondary fluid inlet to the tank, and t_s is the temperature at which the storage medium melts or freezes (e.g., 32°F [0°C]).

The minimum charging rate is actually the negative of the maximum discharging rate and can be given by

$$u_{k,min} = \varepsilon_{d,max} \dot{m}_{f,max} c_f (t_s - t_{f,i}) \qquad (21)$$

where $\varepsilon_{d,k,max}$ is the heat transfer effectiveness for discharging at the current state of storage if the secondary fluid flow rate were at its maximum value of $m_{f,max}$.

In general, the heat transfer effectiveness for charging and discharging at the design flow can be correlated as a function of state of charge using manufacturers' data (e.g., Drees and Braun 1995).

A model for the total building power is also needed. At any time

$$P = P_{noncooling} + P_{plant} + P_{dist} \qquad (22)$$

where $P_{noncooling}$ is the building electrical use that is not associated with the cooling system (e.g., lights), P_{plant} is the power needed to operate the cooling plant, and P_{dist} is the power associated with the distribution of secondary fluid and air through the cooling coils. The models used by Henze et al. (1997a) predict cooling plant and distribution system power with a component-based simulation that is appropriate for simulation studies. Alternatively, for online optimization, plant and distribution system power can be represented with empirical correlations. Drees (1994) used curve-fits of plant power consumption in terms of cooling load and ambient wet-bulb temperature. At any time, the cooling requirement for the chiller is the difference between the building load requirement and the storage discharge rate. The chiller supply temperature is then determined from an energy balance on the chiller and used to evaluate the limits on the storage charging and discharging rates in Equations (21) and (22). The chiller cooling rate must be greater than a minimum value for safe operation and less than the chiller capacity. Drees (1994) correlated the maximum cooling capacity as a function of the ambient wet-bulb temperature and the chiller supply temperature. For simulation studies, a building model may be used to estimate building cooling loads. For online optimization, a forecaster would provide estimates of future building cooling loads.

Dynamic Optimization Methods for Cooling Systems with Thermally Activated Building Systems

Optimal supervisory control for precooling building mass usually requires transient thermal response models of the conditioned zones, because the state of charge cannot be measured and the fully charged and discharged states are not well defined . For passive TES, Equation (12) is retained as the objective function and Equation (22) also applies, with P_{plant} a function of capacity u_k and conditions that affect system COP. One of these conditions is chilled-water return temperature, which is in turn a function of current and past capacities and zone temperatures and past return temperatures: essentially, a conduction transfer function. The constraint on plant capacity [Equation (13)] is retained as well.

Equation (14) is replaced by a transient thermal response model that expresses room temperature x_k in terms of current and past weather, direct heating and cooling (including direct solar and internal gains), and control actions $u_k \ldots u_{k-n}$, and in terms of past room temperatures $x_{k-1} \ldots x_{k-n}$. The comprehensive room transfer function (CRTF) of Seem (1989a) is such a model. Methods and results of training a CRTF have been reported by Armstrong (2000, 2006) and Gayeski et al. (2010b). Equation (15) serves to constrain room temperature during occupied hours and constraint (16) is not normally needed if timestep 0 is assigned to an occupied hour. Note that for TABS the supervisory control cannot generally be expressed in terms of zone temperature set points.

Equation (12), may involve only the chiller power used to meet sensible load, $P = u/\text{COP}(u,t)$. It is possible, with, e.g., liquid desiccant storage or ice storage, to shift latent load. However with enthalpy recovery (see Chapter 44 of the 2008 *ASHRAE Handbook—HVAC Systems and Equipment*), latent loads are a small fraction of total load and not usually considered to be attractive for peak shifting; in this case, DOAS system power may be treated as part of the non-cooling power term in Equation (22). Setting $E = 1$ and $D = 0$ in Equation (12) will produce the highest energy savings.

CONTROL STRATEGIES AND OPTIMIZATION FOR COOLING SYSTEMS

CONTROL STRATEGIES FOR COOLING TOWER FANS

Figure 24 shows a schematic of the condenser loop for a typical chilled-water unit consisting of centrifugal chillers, cooling towers, and condenser water pumps. Typically, the condenser water pump control is directed by the chiller control to provide relatively constant flow for individual chillers. However, the cooling tower cells may be independently controlled to maximize system efficiency.

Typically, cooling tower fans are controlled using a feedback controller that attempts to maintain a temperature set point for the water supplied to the chiller condensers. Often, the condenser water supply temperature set point is held constant. However, a better strategy is to maintain a constant temperature difference between the condenser water supply and the ambient wet bulb (constant approach). Additional savings are possible through optimal control.

With a single feedback controller, the controller output signal must be converted to a specific fan sequence that depends on the number of operating cells and the individual fan speeds. Typically, with the discrete control associated with one- or two-speed tower fans, the set point cannot be realized, resulting in the potential for oscillating tower fan control. Fan cycling can be reduced by using deadbands, "sluggish" control parameters, and/or lower limits for on and off periods.

Braun and Diderrich (1990) demonstrated that feedback control for cooling tower fans could be eliminated by using an open-loop supervisory control strategy. This strategy requires only measuring chiller loading to specify the control and is inherently stable. The tower fan control is separated into two parts: tower sequencing and optimal airflow. For a given total tower airflow, general rules for optimal tower sequencing are used to specify the number of operating cells and fan speeds that give the minimum power consumption for both the chillers and tower fans. The optimal tower airflow is estimated with an open-loop control equation that uses design information for the cooling tower and chiller. The computational procedure is presented in this section, and the control strategy is summarized in a set of steps and sample calculations. The procedures can be adapted to applications where fan and cell performance characteristics vary among cells.

Near-Optimal Tower Fan Sequencing

For variable-speed fans, minimum power consumption results when all cooling tower cells are operated under all conditions. Tower airflow varies almost linearly with fan speed, whereas the fan power consumption varies approximately with the cube of the speed. Thus, for the same total airflow, operating more cells in parallel allows for lower individual fan speeds and lower overall fan power consumption. An additional benefit associated with full-cell operation is lower water pressure drops across the spray nozzles, which results in lower pumping power requirements. However, at very low pressure drops, inadequate spray distribution may adversely affect the thermal performance of the cooling tower.
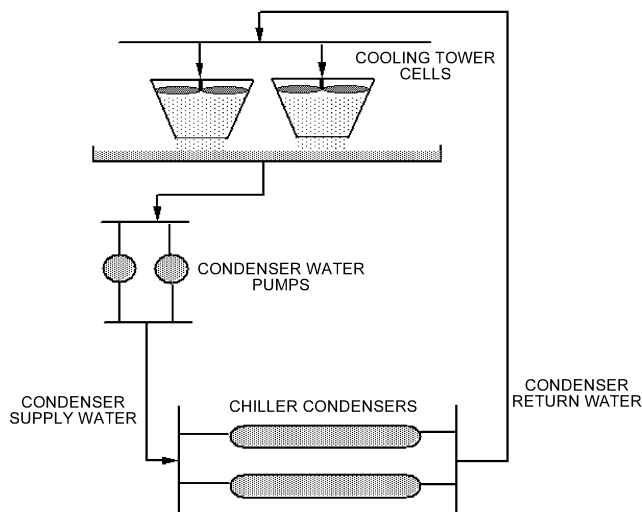

Fig. 24   Condenser Water Loop Schematic

Most cooling towers use multiple-speed rather than continuously adjustable variable-speed fans. In this case, it is not optimal to operate all tower cells under all conditions. The optimal number of cells operating and individual fan speeds depend on the system characteristics and ambient conditions. However, simple relationships exist for the best sequencing of cooling tower fans as capacity is added or removed. When additional tower capacity is required, Braun et al. (1989b) showed that, in almost all practical cases, the speed of the tower fan operating at the lowest speed (including fans that are off) should be increased first. The rules for bringing cell fans online are as follows:

Sequencing Rules

• **All variable-speed fans:** Operate all cells with fans at equal speeds.

DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*

- **Multiple-speed fans:** Activate lowest-speed fans first when adding tower capacity. Reverse for removing capacity.
- **Variable/multiple-speed fans:** Operate all cells with variable fans at equal speeds. Activate lowest-speed fans first when adding tower capacity with multiple-speed fans. Add multiple-speed fan capacity when variable-speed fan speeds match the fan speed associated with the next multiple-speed fan increment to be added.

Similarly, for removing tower capacity, the highest fan speeds are the first to be reduced and sequences defined here are reversed.

These guidelines were derived by evaluating incremental power changes for fan sequencing. For two-speed fans, the incremental power increase associated with adding a low-speed fan is less than that for increasing one to high speed if the low speed is less than 79% of the high fan speed. In addition, if the low speed is greater than 50% of the high speed, then the incremental increase in airflow is greater (and therefore thermal performance is better) for adding the low-speed fan. Most commonly, the low speed of a two-speed cooling tower fan is between one-half and three-quarters of full speed. In this case, tower cells should be brought online at low speed before any operating cells are set to high speed. Similarly, the fan speeds should be reduced to low speed before any cells are brought offline.

For three-speed fans, low speed is typically greater than or equal to one-third of full speed, and the difference between the high and intermediate speeds is equal to the difference between the intermediate and low speeds. In this situation, the best sequencing strategy is to activate the lowest fan speeds first when adding tower capacity and deactivate the highest fan speeds first when removing capacity. Typical three-speed combinations that satisfy these criteria are (1) one-third, two-thirds, and full speed or (2) one-half, three-quarters, and full speed.

Another issue related to control of multiple cooling tower cells with multiple-speed fans concerns the distribution of water flow to the individual cells. Typically, water flow is divided equally among the operating cells. Even though the overall thermal performance of the cooling tower is best when the flow is divided such that the ratio of water-to-airflow rates is identical for all cooling tower cells, equal water flow distribution results in near-optimal performance.
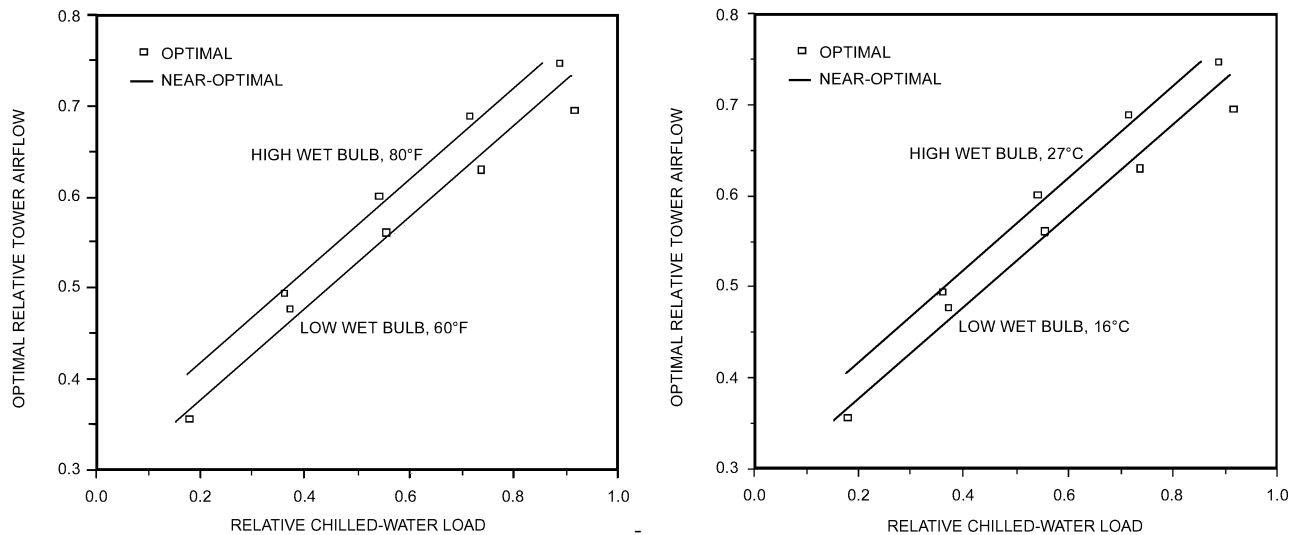
Near-Optimal Tower Airflow

Figure 25 illustrates the tradeoff between the chiller and cooling tower fan power associated with increasing tower airflow for variable-speed fans. As airflow increases, fan power increases with a cubic relationship. At the same time, there is a reduction in the temperature of the water supplied to the condenser of the chiller, resulting in lower chiller power consumption. The minimum total power occurs at a point where the rate of increase in fan power with airflow is equal to the rate of decrease in chiller power. Near the optimum, the total power consumption is not very sensitive to the control. This "flat" optimum indicates extreme accuracy is not needed to determine the optimum control. In general, it is better to have too high rather than too low a fan speed.
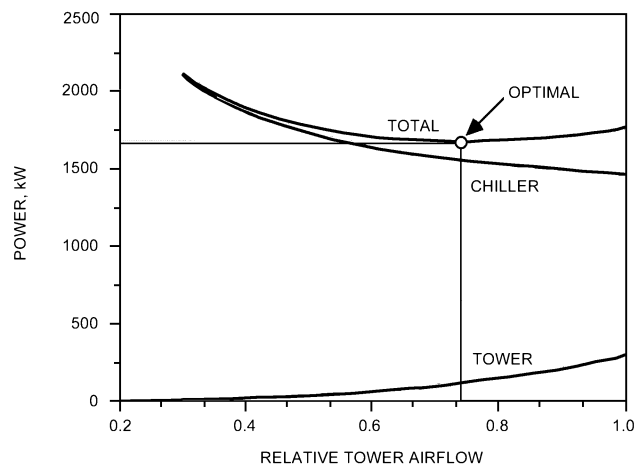

Fig. 25   Tradeoffs Between Chiller Power and Fan Power with Tower Airflow

Braun et al. (1989b) showed that the tower control that minimizes the instantaneous power consumption of a cooling plant varies as a near-linear function of the load over a wide range of conditions. Although optimal control depends on the ambient wet-bulb temperature, this dependence is small compared to the load effect. Figure 26 shows an example of how the optimal tower control varies for a specific plant. The tower airflow as a fraction of the design capacity is plotted as a function of load relative to design load for two different wet-bulb temperatures. For a 20°F (10 K) change in wet-bulb temperature, the optimal control varies only about 5% of the tower capacity. This difference in control results in less than a 1% difference in the plant power consumption. Figure 26 also shows that linear functions work well in correlating the optimal control over a wide range of loads for the two wet-bulb temperatures. Given the insensitivity to wet-bulb temperature and the fact that the load is highly correlated with wet bulb, a single linear relationship is adequate in correlating the optimal tower control in terms of load.

DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*

Fig. 26   Example of Optimal Tower Fan Control

Figure 27 depicts the general form to determine tower airflow as a function of load. The (unconstrained) relative tower airflow is computed as a linear function of the part-load ratio as
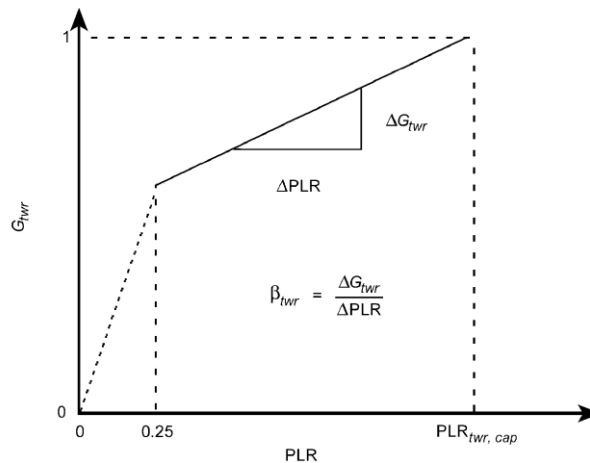
Fig. 27   Fractional Tower Airflow Versus Part-Load Ratio

$$G_{twr} = 1 - \beta_{twr}(\text{PLR}_{twr,cap} - \text{PLR}) \text{ for } 0.25 < \text{PLR} < 1.00 \qquad (23)$$

where

     G_{twr} = tower airflow divided by maximum airflow with all cells operating at high speed
     PLR = chilled-water load divided by design total chiller plant cooling capacity (part-load ratio)
$\text{PLR}_{twr,cap}$ = part-load ratio (value of PLR) at which tower operates at its capacity ($G_{twr} = 1$)
     β_{twr} = slope of relative tower airflow (G_{twr}) versus part-load ratio (PLR) function

The linear relationship between airflow and load is only valid for loads greater than about 25% of the design load. For many installations, chillers do not operate at these small loads. However, for those situations in which chiller operation is necessary below 25% of full load, the tower airflow should be ramped to zero as the load goes to zero according to

$$G_{twr} = 4\text{PLR}[1 - \beta_{twr}(\text{PLR}_{twr,cap} - 0.25)] \text{ for PLR} < 0.25 \qquad (24)$$

The results of either Equation (23) or (24) must be constrained between 0 and 1. This fraction of tower capacity is then converted to a tower control using the sequencing rules of the section on Near-Optimal Tower Fan Sequencing.

The variables of the open-loop linear control Equation (23) that yield near-optimal control depend on the system's characteristics. Detailed measurements may be taken over a range of conditions and used to accurately estimate these variables. However, this requires measuring component power consumption along with considerable time and expertise, and may not be cost effective unless performed by on-site plant personnel. Alternatively, simple estimates of these parameters may be obtained using design data.

Open-Loop Parameter Estimates Using Design Data. Good estimates for the parameters of Equation (23) may be determined analytically using design information as summarized in Table 4. These estimates were derived by Braun and Diderrich (1990) by applying optimization theory to a simplified mathematical model of the chiller and cooling tower, assuming that the tower fans are sequenced in a near-optimal manner. In general, these estimates are conservative in that they should provide greater rather than less than the optimal tower airflow. The results given in Table 4 for variable-speed fans should also provide adequate estimates for three-speed fans.

Table 4   Parameter Estimates for Near-Optimal Tower Control Equation

Parameter	One-Speed Fans	Two-Speed Fans	Variable-Speed Fans
$\mathrm{PLR}_{twr,cap}$	PLR_0	$\sqrt{2}\mathrm{PLR}_0$	$\sqrt{3}\mathrm{PLR}_0$
$\square_{twr}$	$\dfrac{1}{\mathrm{PLR}_{twr,cap}}$	$\dfrac{2}{3\mathrm{PLR}_{twr,cap}}$	$\dfrac{1}{2\mathrm{PLR}_{twr,cap}}$

Note:

$$\mathrm{PLR}_0 = \frac{1}{\sqrt{\dfrac{P_{ch,des}}{P_{twr,des}}}\, S_{cwr,des}(a_{twr,des}+r_{twr,des})}$$

Design factors that affect the parameter estimates given in Table 4 are the (1) ratio of chiller power to cooling tower fan power at design conditions $P_{ch,des}/P_{twr,des}$, (2) sensitivity of chiller power to changes in condenser water return temperature at design conditions $S_{cwr,des}$, and (3) sum of the tower approach and range at design conditions ($a_{twr,des} + r_{twr,des}$). Chiller power consumption at design conditions is the total power consumption of all plant chillers operating at their design cooling capacity. Likewise, the design tower fan power is the total power associated with all tower cells operating at high speed. As the ratio of chiller power to tower fan power increases, it becomes more beneficial to operate the tower at higher airflows. This is reflected in a decrease in the part-load ratio at which the tower reaches its capacity, $\mathrm{PLR}_{twr,cap}$. If the tower airflow were free (i.e., zero fan power), then $\mathrm{PLR}_{twr,cap}$ would go to zero, and the best strategy would be to operate the towers at full capacity independent of the load. A typical value for the ratio of the chiller power to the cooling tower fan power at design conditions is 10.

The chiller sensitivity factor $S_{cwr,des}$ is the incremental increase in chiller power for each degree increase in condenser water temperature as a fraction of the power or

$$S_{cwr,des} = \frac{\text{Change in chiller power}}{\text{Change in cond. water return temp.} \times \text{Chiller power}} \tag{25}$$

If the chiller power increases by 2% for a 1°F (0.5 K) increase in condenser water temperature, $S_{cwr,des}$ is equal to 0.02/°F (0.04/K). A large sensitivity factor means that the chiller power is very sensitive to the cooling tower control favoring operation at higher airflow rates (low $\mathrm{PLR}_{twr,cap}$). The sensitivity factor should be evaluated at design conditions using chiller performance data. Typically, the sensitivity factor is between 0.01 and 0.03/°F (0.02 and 0.06/K). For multiple chillers with different performance characteristics, the sensitivity factor at design conditions may be estimated by

$$S_{cwr,des} = \frac{\sum_{i=1}^{N_{ch}} S_{cwr,des,i}\, P_{ch,des,i}}{\sum_{i=1}^{N_{ch}} P_{ch,des,i}} \tag{26}$$

where $S_{cwr,des,i}$ is the sensitivity factor and $P_{ch,des,i}$ is the power consumption for the ith chiller at the design conditions, and N_{ch} is the total number of chillers.

The design approach to wet bulb $a_{twr,des}$ is the temperature difference between the condenser water supply and the ambient wet bulb for the tower, operating at its air and water flow capacity at plant design conditions. The design range $r_{twr,des}$ is the water temperature difference across the tower at these same conditions (condenser water return minus supply temperature). The sum of $a_{twr,des}$ and $r_{twr,des}$ is the temperature difference between the tower inlet and the ambient wet bulb and represents a measure of the tower's capability to reject heat to ambient relative to the system requirements. A small temperature difference (tower approach plus range) results from a high tower heat transfer effectiveness or high water flow rate and yields lower condenser water temperatures with lower chiller power consumption. Typical values for the design approach and range are 7 and 10°F (4 and 6 K).

The part-load ratio associated with the tower operating at full capacity, $\mathrm{PLR}_{twr,cap}$ may be greater than or less than one. Values less than unity imply that, from an energy point of view, the tower is not sized for optimal operation at design load conditions and that it should operate at its capacity for a range of loads less than the design load. Values of $\mathrm{PLR}_{twr,cap}$ greater than one imply that the tower is oversized for the design load and that it should never operate at its capacity.

For multiple chillers with very different performance characteristics, different open-loop parameters may be used for any combination of operating chillers. The sensitivity factors and chiller design power used to determine the open-loop control parameters in Table 4

should be estimated for each combination of operating chillers, and the part-load ratio used in Equation (23) should be determined using the design capacity for the operating chillers (not all chillers). In this case, N_{ch} in Equation (26) represents the number of operating chillers.

Open-Loop Parameter Estimates Using Plant Measurements. Energy consumption may be somewhat further reduced by determining the open-loop control parameters from plant measurements. However, this results in additional complexity associated with implementation. One method for estimating the open-loop control parameters of Equation (23) from plant measurements involves performing a set of one-time trial-and-error experiments. At a given set of conditions (i.e., cooling load and ambient conditions), the optimal tower control is estimated by varying the fan settings and monitoring the total chiller and fan power consumption. Each tower control setting and load condition must be maintained for a sufficient time for the power consumption to approach steady-state and to hold the chilled-water supply temperature constant. The control setting that produces the minimum total power consumption is deemed optimal. This set of experiments is performed for a number of chilled-water cooling loads and the best-fit straight line through the resulting data points is used to estimate the parameters of Equation (23). As initial control settings for each load, Equation (23) may be used with estimates from design data as summarized in the previous section.

Another method for estimating the variables of Equation (23) uses an empirical model for total power consumption that is fit to plant measurements. The control that minimizes the power consumption associated with the model is then determined analytically. The section on Control Optimization Methods describes a general method for determining linear control relations in this manner using a quadratic model. For cooling tower fan control, chiller and fan power consumption are correlated with load and tower airflow for a constant chilled-water supply temperature using a quadratic function as follows:

$$P = a_0 + a_1 \text{PLR} + a_2 \text{PLR}^2 + a_3 G_{twr}$$
$$+ a_4 G^2_{twr} + a_5 \text{PLR} \times G_{twr} \tag{27}$$

where a_0 to a_5 are empirical constants determined through linear regression applied to measurements. For the quadratic function of Equation (27), the tower airflow that results in minimum power is a linear function of the PLR. The parameters of the open-loop control Equation (23) are then

$$\text{PLR}_{twr,cap} = - \frac{a_3 + 2a_4}{a_5} \tag{28}$$

$$\beta_{twr} = - \frac{a_5}{2a_4} \tag{29}$$

For multiple chillers with very different performance characteristics, different open-loop parameters can be determined for any combination of operating chillers. In this case, separate correlations for near-optimal airflow or power consumption must be determined for each chiller combination.

Overrides for Equipment Constraints

The fractional tower airflow as determined by Equations (23) or (24) must be bounded between 0 and 1 according to the physical constraints of the equipment. Additional constraints on the temperature of the supply water to the chiller condensers are necessary to avoid potential chiller maintenance problems. Many (older) chillers have a low limit on the condenser water supply temperature that is necessary to avoid lubrication migration from the compressor. A high-temperature limit is also necessary to avoid excessively high pressures in the condenser, which can lead to compressor surge in centrifugal chillers. If condenser water temperature falls below the low limit, then it is necessary to override the open-loop tower control and reduce tower airflow to go above this limit. Similarly, if the high limit is exceeded, then tower airflow should be increased as required.

Implementation

Before commissioning, the parameters of the open-loop control Equation (23) must be specified. These parameters are estimated using Table 4. After the system is in operation, these parameters may be fine-tuned with measurements as outlined previously. If multiple chillers have significantly different performance characteristics, it may be advantageous to determine different parameters for Equation (23) depending on the combination of operating chillers.

The relative tower airflow must be converted to a specific set of tower fan settings using the sequencing rules defined previously. This involves defining a relationship (i.e., table) for fan settings as a function of tower airflow. The table is constructed by defining the best fan settings for each possible increment of airflow. The conversion process between the continuous output of Equations (23) or (24) and the fan control involves choosing the set of discrete fan settings from the table that produces a tower airflow closest to the desired flow. However, in general, it is better to have greater rather than less than the optimal airflow. A good general rule is to choose the set of discrete fan controls that results in a relative airflow that is closest to, but not more than 10% less than, the output of Equations (23) or (24).

With the parameters of Equation (23) specified, the following procedure is applied at each decision interval (e.g., 15 min) to determine the tower control:

1. If the temperature of supply water to the chiller condenser is less than the low limit, then reduce tower airflow by one increment according to the near-optimal sequencing rules and exit the algorithm. Otherwise go to step 2.
2. If the temperature of supply water to the chiller condenser is greater than the high limit, then increase tower airflow by one increment according to the near-optimal sequencing rules and exit the algorithm. Otherwise go to step 3.
3. Determine the chilled-water load relative to the design load.

4. If the chilled-water load has changed by a significant amount (e.g., 10%) since the last control change, then go to step 5. Otherwise, exit the algorithm.
5. If the part-load ratio is greater than 0.25, then compute the near-optimal tower airflow as a fraction of tower capacity G_{twr} with Equation (23). Otherwise, determine G_{twr} with Equation (24).
6. Limit G_{twr} to keep the change from the previous decision interval less than a minimum value (e.g., less than 0.1 change).
7. Restrict the value of G_{twr} between 0 and 1.
8. Convert the value of G_{twr} to a specific set of control functions for each of the tower cell fans according to the near-optimal sequencing rules.

This procedure requires some estimate of the chilled-water load, along with a measurement of the condenser water supply temperature. However, the accuracy of the load estimates is not extremely critical. In general, near-optimal control determined with load estimates that are accurate to within 5 to 10% results in total power consumption that is within 1% of the minimum. The best method for determining the chilled-water load is from the product of the measured chilled-water flow rate and the temperature difference between the chilled-water return and supply. For systems that use constant flow pumping to the chillers, the flow rates may be estimated from design data for the pumps and system pressure-drop characteristics.

Example 1. Consider an example plant consisting of four 550 ton (2000 kW) chillers with four cooling tower cells, each having two-speed fans. Each chiller consumes approximately 330 kW (340 kW) at the design capacity, and each tower fan uses 40 kW (42 kW) at high speed. At design conditions, the chiller power increases approximately 6.6 kW (12.2 kW) for a 1°F (1 K) increase in condenser water temperature, giving a sensitivity factor of 6.6/330 or 0.02/°F (12.2/340 or 0.036/K). The tower design approach and range from manufacturer's data are 7 and 10°F (4 and 6 K).

Solution:

The first step in applying the open-loop control algorithm to this problem is determining the parameters of Equation (23) from the design data. From Table 4, the part-load ratio at which operation of the tower is at its capacity is estimated for the two-speed fans as

$$\text{PLR}_{twr,cap} = \cfrac{1}{\sqrt{\cfrac{1}{2}(0.02/°F)\cfrac{4 \times 330 \text{ kW}}{4 \times 40 \text{ kW}}(7 + 10°F)}} = 0.84 \qquad \left(\text{PLR}_{twr,cap} = \cfrac{1}{\sqrt{\cfrac{1}{2}(0.036/K)\cfrac{4 \times 340 \text{ kW}}{4 \times 42 \text{ kW}}(4 + 6)}} = 0.83\right)$$

and the slope of the fractional airflow versus part-load ratio is estimated to be

$$\beta_{twr} = \cfrac{2}{3 \times 0.84} = 0.79 \qquad \left(\beta_{twr} = \cfrac{2}{3 \times 0.83} = 0.80\right)$$

Given these parameters and the part-load ratio, the fractional tower airflow is estimated as

IF (PLR > 0.25) THEN
$\qquad G_{twr} = 1 - \beta_{twr}(\text{PLR}_{twr,cap} - \text{PLR})$
ELSE
$\qquad G_{twr} = 4\text{PLR}[1 - \beta_{twr}(\text{PLR}_{twr,cap} - 0.25)]$
$\qquad G_{twr} = \text{MIN}[1, \text{MAX}(0, G_{twr})]$

To convert G_{twr} into a specific tower control, the tower sequencing must be defined. The following table gives this information in a form that specifies the relationship between G_{twr} and tower control for this example.

		Cooling Tower Fan Sequencing for Example 1			
		Tower Fan Speeds			
Sequence No.	G_{twr}	**Cell #1**	**Cell #2**	**Cell #3**	**Cell #4**
1	0.125	Low	Off	Off	Off
2	0.250	Low	Low	Off	Off
3	0.375	Low	Low	Low	Off
4	0.500	Low	Low	Low	Low
5	0.625	High	Low	Low	Low
6	0.750	High	High	Low	Low
7	0.875	High	High	High	Low
8	1.000	High	High	High	High

For a specific chilled-water load, the fan control should be the sequence of tower fan settings from the table that results in a value of G_{twr} that is closest to, but not more than 10% less than, the output of Equations (23) or (24). Note that this example assumes that proper water flow can be maintained over all cooling tower cells.

CHILLED-WATER RESET WITH FIXED-SPEED PUMPING

Figure 28 shows a common configuration using fixed-speed chilled-water pumps with two-way valves at the cooling coils. A two-way bypass valve controlled to maintain a fixed pressure difference between the main supply and return lines is used to ensure relatively constant flow through chiller evaporators and reduce pressure drop and pumping costs at low loads. However, additional pump and chiller power savings can be realized by adjusting the chilled-water supply temperature to keep some cooling coil valves open and thereby minimize the bypass flow.

Ideally, the chilled-water temperature should be adjusted to maintain all discharge air temperatures with a minimal number of cooling-coil control valves in a saturated (fully open) condition. The procedure described in this section is designed to accomplish this goal in a reliable and stable manner that reacts quickly to changing conditions.
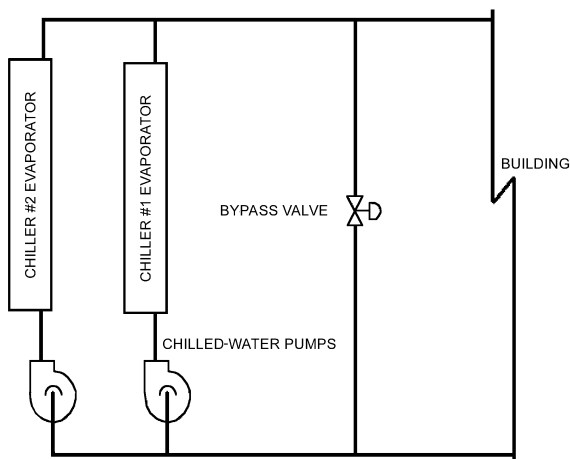

Fig. 28   Typical Chilled-Water Distribution for Fixed-Speed Pumping

Pump Sequencing

In plants where one chilled-water pump is dedicated to each chiller, the sequencing of chilled-water pumps is defined by the sequencing of chillers. In some installations, the chiller pumps are not dedicated to chillers, but instead are arranged in parallel, sharing common headers. In this case, the order for bringing pumps online and offline and the conditions for adding or removing chilled-water pump capacity must be specified. For pumps of different capacities, the logical order for bringing pumps online is from small to large. For pumps of similar capacity, the most efficient pumps should be brought online first and taken offline last.

Optimal Chilled-Water Temperature

One method for determining the optimal chilled-water temperature is to monitor the water control valve positions of "representative" air handlers and to adjust the set temperature incrementally at fixed decision intervals until a single control valve is fully open. The representative air handlers should be chosen to include load diversity at all times and ensure reliable data. One difficulty of this control approach is that valve position data are often unreliable. The valve could be stuck open or the saturation indicator could be faulty. This problem can be overcome by also monitoring discharge air temperatures, using them as a consistency check on valve position data. If a valve is unsaturated, this implies that the coil has sufficient capacity to maintain the discharge air temperature near the set point. Conversely, if a valve remains saturated at 100% open, the discharge air temperature should ultimately increase above the set point. These considerations lead to the following simple rules for increasing or decreasing the chilled-water set point in response to valve position and discharge air temperature data.

• If all water valves are unsaturated or the discharge air temperatures associated with all saturated valves are lower than the set point, increase the chilled-water set temperature.
• If more than one valve is saturated at 100% open and their corresponding discharge air temperatures are greater than their set points, decrease the chilled-water temperature.

In implementing these rules, a fixed increment for increasing or decreasing the chilled-water temperature must be chosen. A small increment results in more stable control, but also results in a slow response to sudden changes in load or supply air temperature set points. Using a first-order approximation, the chilled-water temperature can be reset in response to sudden changes in load and supply air temperature set point according to

$$t_{chws} = t_{as} - \frac{PLR}{PLR_o}(t_{as,o} - t_{chws,o}) \qquad (30)$$

where

   t_{chws} = new chilled-water set-point temperature
     t_{as} = current supply air set-point temperature
   PLR = current part-load ratio (chiller load divided by total design load for all chillers)
 $t_{chws,o}$ = chilled-water set point associated with last control decision
   $t_{as,o}$ = supply air set point associated with last control decision
  PLR_o = part-load ratio associated with last control decision

Equation (30) assumes that the chilled-water temperature associated with the last control decision was optimal. As a result, it only applies to anticipating the effects of significant changes in the load and supply air set-point temperature on the optimal chilled-water set point. The "bump-and-wait" strategy fine-tunes the chilled-water supply temperature when the load and supply air set point are stable.

For a variable-air-volume system, the supply air set points are most often constant and identical for all air-handling units. However, for a constant-air-volume system, these set points may vary with different air handlers. In this case, the supply air set point to use in Equation (30) should be a capacity-weighted average value for the representative air handlers.

Equation (30) indicates that the optimal chilled-water supply temperature increases with increasing supply air temperature and decreasing load. This is because these changes cause the cooling-coil valves to close; optimal control involves keeping at least one valve open. Increasing supply air temperature causes the cooling-coil valves to close somewhat because of a larger average temperature difference for heat transfer between the water and air. A lower load requires smaller air-to-water temperature differences, which also leads to control valves closing.

Overrides for Equipment and Comfort Constraints

For a given chiller load, the chilled-water temperature has both upper and lower limits. The lower limit is necessary to avoid ice formation on the evaporator tubes of the chiller. This limit depends primarily on the load in relation to the size of the evaporator or, in other words, the temperature difference between the chilled water and refrigerant. At small temperature differences (large area or small load), the evaporator can tolerate a lower chilled-water temperature to avoid freezing than at large temperature differences. The lower limit on the chilled-water set point should be evaluated at the design load, because the overall system performance is improved by increasing chilled-water temperature above this limit for loads less than design. This lower limit can range from 38 to 44°F (3 to 7°C) when chilled water is used for dehumidification, to 55°F (13°C) for TABS, and as high as 60°F (15.6°C) for chilled water serving radiant or chilled-beam cooling systems.

An upper limit on the chilled-water temperature arises from comfort constraints associated with the zones and the possibility of microbial growth associated with high humidities. For the available flows, the chilled-water temperature should be low enough to provide discharge air at a temperature and humidity sufficient to maintain all zones in the comfort region and avoid microbial growth. This upper limit varies with both load and entering air conditions and is accounted for by monitoring the zone conditions to ensure that they are in the comfort zone. If zone temperatures or humidities are not within reasonable bounds, then the discharge air temperature set point should be lowered. For radiant panel, chilled-beam, or TAB systems, there is no specific upper chilled-water limit as long as the load is satisfied.

Implementation

At each decision interval (e.g., 5 min), the following algorithm would be applied for determining the optimal chilled-water set-point temperature:

1. Determine the time-averaged total chilled-water load for the previous decision interval.
2. If the chilled-water load or supply air set-point temperature has changed by a significant amount (e.g., 10%) since the last control change, then estimate a new optimal chilled-water set point with Equation (30) and go to step 6. Otherwise, go to step 3.
3. Determine the time-averaged position of (or controller output for) the cooling-coil water valves and corresponding discharge air temperatures for representative air handlers.
4. If more than one valve is saturated at 100% open and their corresponding supply air temperatures are greater than set point (e.g., 1°F [0.5 K]), then decrease the chilled-water temperature by a fixed amount (e.g., 0.5°F [0.25 K]) and go to step 6. Otherwise, go to step 5.
5. If all water valves are unsaturated or the supply air temperatures associated with all valves that are saturated are lower than the set point, then raise the chilled-water set temperature by a fixed amount (e.g., 0.5°F [0.25 K]). Otherwise, exit the algorithm with the chilled-water set point unchanged.
6. Limit the chilled-water set-point temperature between the upper and lower limits dictated by comfort, humidity, and equipment safety.

Implementing this algorithm requires some estimate of the chilled-water load, along with a measurement of the discharge air temperatures and control valve positions. However, a highly accurate estimate of the load is not necessary.

CHILLED-WATER RESET WITH VARIABLE-SPEED PUMPING

Figure 29 shows a common configuration for systems using variable-speed chilled-water pumps with primary/secondary water loops. The primary pumps are fixed speed and are generally sequenced with chillers to provide a relatively constant flow of water through the chiller evaporators. The secondary chilled-water pumps are variable speed and are typically controlled to maintain a specified set point for pressure difference between supply and return flows for the cooling coils.

DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*
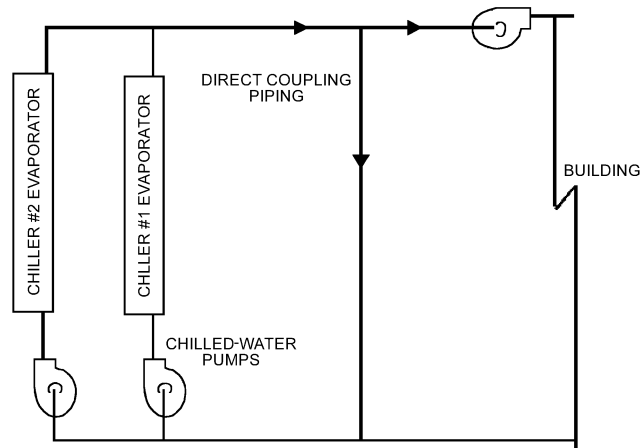
**Fig. 29   Typical Chilled-Water Distribution for Primary/Secondary Pumping**

Although variable-speed pumps are usually used with primary/secondary chilled-water loops, they may also be applied to systems with a single chilled-water loop. In either case, variable-speed pumps offer the potential for a significant operating cost saving when both chilled-water and pressure differential set points are optimized in response to changing loads. This section presents an algorithm for determining near-optimal values of these control variables.

Optimal Differential Pressure Set Points

In practically all variable-speed chilled-water pumping applications, pump speed is controlled to maintain a constant pressure differential between the main chilled-water supply and return lines. However, this approach is not optimal. To maintain a constant pressure differential with changing flow, the control valves for the air-handling units must close as the load (i.e., flow) is reduced, resulting in an increase in the flow resistance. The best strategy for a given chilled-water set point is to reset the differential pressure set point to maintain all discharge air temperatures with at least one control valve in a saturated (fully open) condition. This results in a relatively constant flow resistance and greater pump savings at low loads. With variable differential pressure set points, optimizing the chilled-water loop is described in terms of finding the chilled-water temperature that minimizes the sum of chiller and pumping power, with pump control dependent on set point and load.

Near-Optimal Chilled-Water Set Point

The optimal chilled-water supply temperature at a given load results from a tradeoff between chiller and pumping power, as illustrated in Figure 30. As the chilled-water temperature increases, chiller power is reduced due to a reduction in the lift requirements of the chiller. For a higher set temperature, more chilled-water flow is necessary to meet the load requirements, and the pumping power requirements increase. The minimum total power occurs at a point where the rate of increase in pumping power with chilled-water temperature is equal to the rate of decrease in chiller power. This optimal set point moves to lower values as the load increases.
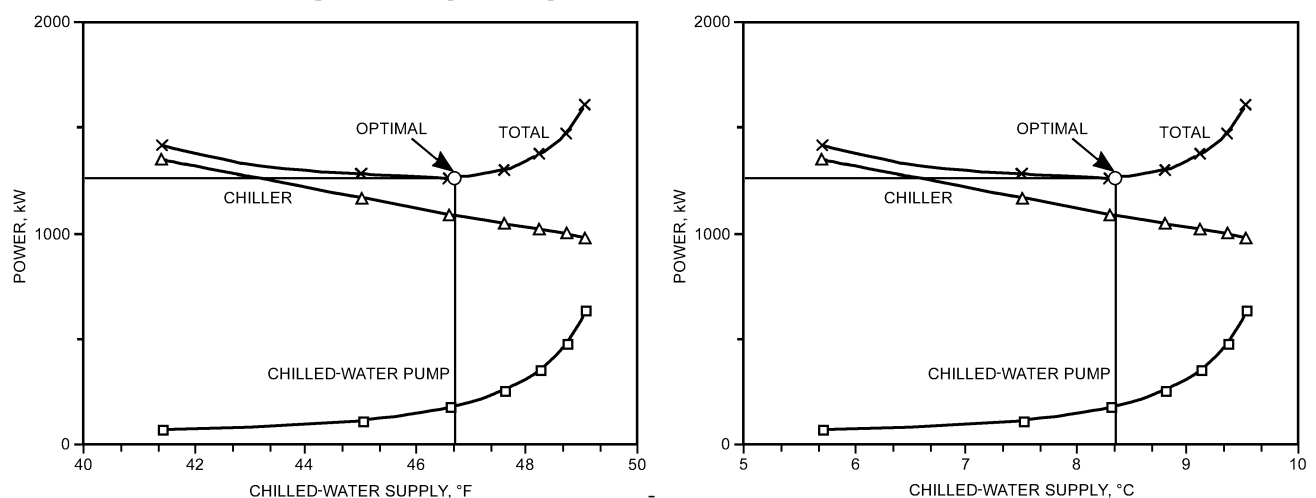
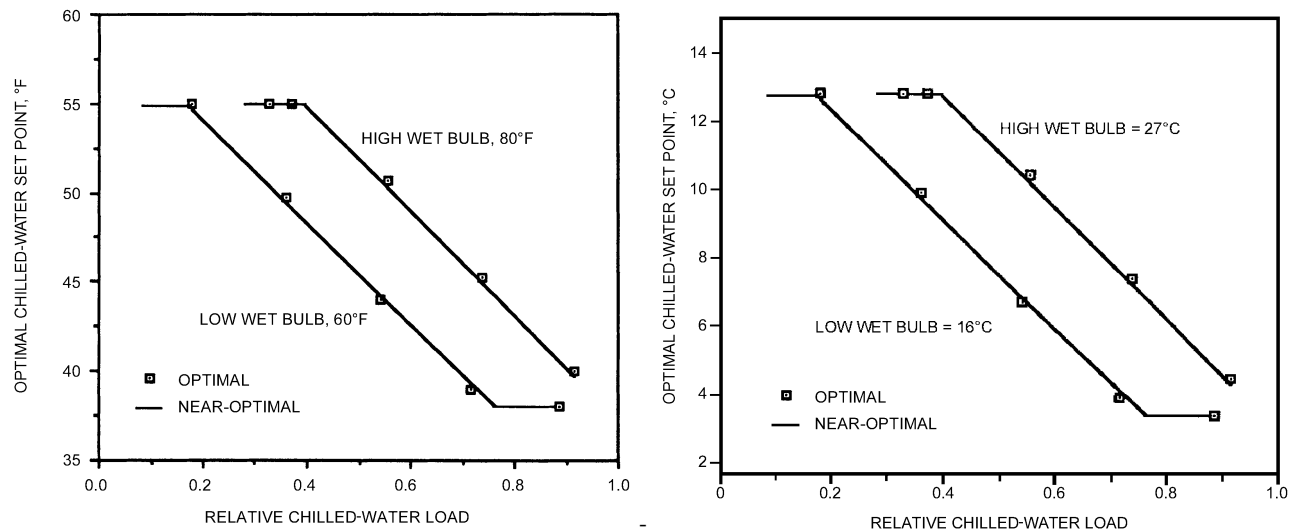**Fig. 30   Tradeoff of Chiller and Pump Power with Chilled-Water Set Point**

Fig. 31   Comparisons of Optimal Chilled-Water Temperature

Braun et al. (1989b) demonstrated that the optimal chilled-water set point varies as a near-linear function of both load and wet-bulb temperature over a wide range of conditions. Figure 31 shows an example of how the optimal set point varies for a specific plant. The set point is plotted as a function of load relative to design load for two different wet-bulb temperatures. In general, the optimal chilled-water temperature decreases with load because pump power becomes a larger fraction of total power. A lower set-point limit is set to avoid conditions that could form ice on evaporator tubes or too high a chiller "lift," and an upper limit is established to ensure adequate cooling-coil dehumidification. For a given load, the chilled-water set point increases with wet-bulb temperature because energy transfer across each cooling coil is proportional to the difference between its entering air wet-bulb temperature and the entering water temperature (the chilled-water set point). For a constant load, this temperature difference is constant and the chilled-water supply temperature increases linearly with entering air wet-bulb temperature.

The results in Figure 31 were obtained for a system where both the chilled-water supply and supply air set points to the zones were optimized. For this case, the supply air temperatures varied between 55°F (13°C) at high loads and 60°F (16°C) at low loads. More typically, supply air temperatures are constant at 55°F (13°C), and the variation in chilled-water supply temperature is smaller than that shown in Figure 30.

Figure 32 depicts the general form for an algorithm to determine chilled-water supply set points as a function of load and the average wet-bulb temperature entering the cooling coils. A normalized difference between the entering air wet-bulb temperature and the chilled-water supply temperature is shown as a linear function of the part-load ratio. The (unconstrained) chilled-water set point is determined as
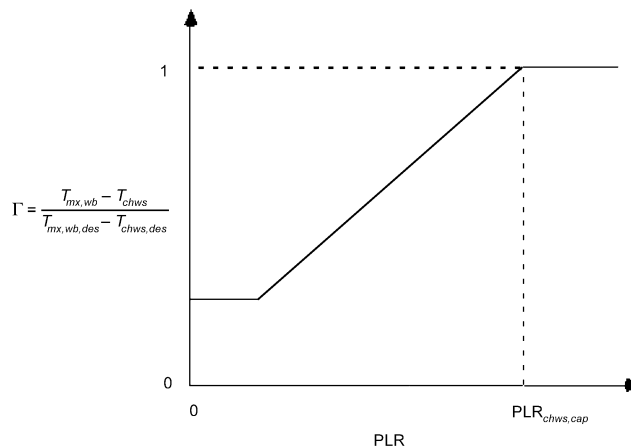

Fig. 32   Dimensionless Chilled-Water Set Point Versus Part-Load Ratio

$$t_{chws} = t_{mx,wb} - \Gamma(t_{mx,wb,des} - t_{chws,des}) \qquad (31)$$

where

$$\Gamma = 1 - \beta_{chws}(PLR_{chws,cap} - PLR) \qquad (32)$$

t_{chws} = chilled-water supply temperature set point

$t_{mx,wb}$ = average or "representative" wet-bulb temperature of air entering cooling coils
$t_{chws,des}$ = chilled-water supply temperature at design conditions
$t_{mx,wb,des}$ = wet-bulb temperature of air entering cooling coils at design conditions
PLR = chilled-water load divided by the total chiller cooling capacity (part-load ratio)
$PLR_{chws,cap}$ = part-load ratio (value of PLR) at which $\Gamma = 1$
β_{chws} = slope of the Γ versus part-load ratio (PLR) function

For RCP and chilled-beam systems that provide sensible cooling only, replace $t_{mx,wb,des}$ and $t_{mx,wb,des}$ by the corresponding zone operative temperatures. The result of Equation (31) must be constrained between upper and lower limits dictated by equipment safety (evaporator freezing), machine operating envelope, and comfort and humidity concerns.

The variables of Equation (32) that yield near-optimal control depend on the system characteristics. Detailed measurements over a range of conditions may be used to estimate these parameters. However, this requires measuring component power consumption along with considerable time and expertise, and may not be cost effective unless performed by on-site plant personnel. Alternatively, simple estimates of these parameters may be obtained using design data.

Table 5 Parameter Estimates for Near-Optimal Chilled-Water Set Point Equation

Parameter	Estimate
$PLR_{chws,cap}$	$\sqrt{\dfrac{1}{3} \times \dfrac{P_{ch,des}}{P_{chwp,des}} S_{chws,des}(t_{mx,wb,des} - t_{chws,des})}$
β_{chws}	$\dfrac{0.5}{PLR_{chws,cap}}$

Open-Loop Parameter Estimates Using Design Data. Reasonable estimates of the parameters of Equation (32) may be determined analytically using design information as summarized in Table 5. These estimates were derived by applying optimization theory to a simplified mathematical model of the chiller and secondary-loop water pumps, assuming that a differential pressure reset strategy is used, pump efficiencies are constant, and the supply air temperature is not varied in response to changes in chilled-water supply temperature. In general, these parameter estimates are conservative in that they should provide a relatively low estimate of the optimal chilled-water set point.

The design factors that affect the parameter estimates given in Table 5 are the (1) ratio of the chiller power to chilled-water pump power at design conditions $P_{ch,des}/P_{chwp,des}$, (2) chiller power's sensitivity to changes in chilled-water temperature at design conditions $S_{chws,des}$, and (3) difference between design entering air wet-bulb temperature to the cooling coil and the chilled-water supply temperature $(t_{mx,wb,des} - t_{chws,des})$.

Chiller power consumption at design conditions is the total power consumption of all plant chillers operating at their design cooling capacity. Likewise, the design pump power is the total power associated with all secondary chilled-water supply pumps operating at high speed. As the ratio of chiller power to pump power increases, it becomes more beneficial to operate the chillers at higher chilled-water temperatures and the pumps at higher flows. This is reflected in an increase in $PLR_{chws,cap}$. If chiller power were free, then $PLR_{chws,cap}$ would go to zero, and the best strategy would be to operate the chillers at the minimum possible set point, resulting in low chilled-water flow rates. Typical values for the ratio of chiller power to pump power at design conditions are between 10 and 20, depending mainly on whether primary/secondary pumping is used.

Chiller sensitivity factor $S_{chws,des}$ is the incremental increase in chiller power for each degree decrease in chilled-water temperature as a fraction of the power:

$$S_{cwr,des} = \frac{\text{Increase in chiller power}}{\text{Decrease in chilled-water temp} \times \text{Chiller power}}$$

(33)

If chiller power increases by 2% for a 1°F (0.5 K) decrease in chilled-water temperature, then $S_{chws,des}$ is equal to 0.02/°F (0.04/K). A large sensitivity factor means that chiller power is very sensitive to the set point control favoring operation at higher set point temperatures and flows (higher $PLR_{chws,cap}$). The sensitivity factor should be evaluated at design conditions using chiller performance data. Typically, the sensitivity factor is between 0.01 and 0.03/°F (0.02 and 0.06/K). For multiple chillers with different performance characteristics, the sensitivity factor at design conditions is estimated as

$$S_{chws,des} = \frac{\sum_{i=1}^{N_{ch}} S_{chws,des,i} P_{ch,des,i}}{\sum_{i=1}^{N_{ch}} P_{ch,des,i}}$$

(34)

DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*

where $S_{chws,des,i}$ is the sensitivity factor and $P_{ch,des,i}$ is the power consumption for the ith chiller at design conditions, and N_{ch} is the total number of chillers.

The design difference between coil inlet air wet-bulb temperature and entering water temperature should be evaluated for a typical air handler operating at design load and flows. A small temperature difference results from a high coil heat transfer effectiveness or high water flow rate, allowing higher chilled-water temperatures with lower chiller power consumption. This is evident from Equation (31), where chilled-water set point decreases linearly with $(t_{mx,wb,des} - t_{chws,des})$ for a given Γ, and Γ is inversely related to the square root of $(t_{mx,wb,des} - t_{chws,des})$. Typically, this temperature difference is about 20°F (10 K).

Example 2. Consider an example plant with primary/secondary chilled-water pumping. There are four 550 ton (2000 kW) chillers, each with a dedicated primary pump. Each chiller consumes approximately 330 kW (340 kW) at design capacity. At design conditions, chiller power increases approximately 6.6 kW (2.2 kW) for a 1°F (0.5 K) decrease in chilled-water temperature, giving a sensitivity factor of 6.6/330 or 0.02/°F (2.2/340 or 0.04/K). The design chilled-water set point is 42°F (5°C), and the coil entering wet-bulb temperature is 62°F (16°C) at design conditions. The secondary loop uses three identical 60 hp (45 kW) chilled-water pumps: one with a variable-speed and two with fixed-speed motors.

Solution:

The first step in applying the open-loop control algorithm to this problem is determining the parameters of Equation (32) from the design data. From Table 5, the part-load ratio at which the chilled-water temperature reaches a minimum (with the design entering wet-bulb temperature to the coils) is

$$\text{PLR}_{chws,cap} = \sqrt{\frac{1}{3} \times \frac{4 \times 330 \text{ kW}}{3 \times 60 \text{ hp} \times 0.75 \text{ kW/hp}}(0.02/°\text{F})(20°\text{F})} = 1.14$$

and the slope of the set point versus part-load ratio is estimated to be

$$\beta_{chws} = \frac{\frac{0.5}{1.14}}{} = 0.44$$

Given these parameters and the part-load ratio, the unconstrained chilled-water set-point temperature from Equations (9) and (10) is then

$$t_{chws} = t_{mx,wb} - [1 - 0.44(1.14 - \text{PLR})]20$$

(The first step in applying the open-loop control algorithm to this problem is determing the parameters of Equation (32) from the design data. From Table 5, the part-load ratio at which the chilled-water temperature reaches a minimum (with the design entering wet-bulb temperature to the coils) is

$$\text{PLR}_{chws,cap} = \sqrt{\frac{1}{3} \times \frac{4 \times 340}{3 \times 45}(0.036/\text{K})(11)} = 1.15$$

and the slope of the set point versus part-load ratio is estimated to be

$$\beta_{chws} = \frac{\frac{0.5}{1.15}}{} = 0.43$$

Given these parameters and the part-load ratio, the unconstrained chilled-water set-point temperature from Equations (9) and (10) is then

$$t_{chws} = t_{mx,wb} - [1 - 0.43(1.15 - \text{PLR})]11)$$

Pump Sequencing

Variable-speed pumps are sometimes used in combination with fixed-speed or other variable-speed pumps. Pump sequencing involves determining both the order and point that pumps should be brought online and offline.

Pumps should be brought online in an order that allows a continuous variation in flow rate and maximized operating efficiency of the pumps at each switch point for the specific pressure loss characteristic. For a combination of fixed- and variable-speed pumps, at least one variable-speed pump should be brought online before any fixed-speed pumps. For single-loop systems (i.e., no secondary loop) with variable-speed pumps, the pressure drop characteristics change when chillers are added or removed and the optimal sequencing of pumps depends on the sequencing of chillers.

An additional pump should be brought online whenever the current set of pumps is operating at full capacity and can no longer satisfy the differential pressure set point. This situation can be detected by monitoring the differential pressure or the controller output signal. Insufficient pump capacity leads to extended periods with differential pressures that are less than the set point and a controller output that is saturated at 100%. A pump may be taken offline whenever the remaining pumps have sufficient capacity to maintain the differential pressure set point. This condition can be determined by comparing the current (time-averaged) controller output with the controller output (time-averaged) at the point just after the last pump was brought online. The pump can be brought offline when the current output is less than the switch point value by a specified dead band (e.g., 5%).

Overrides for Equipment and Comfort Constraints

The chilled-water temperature is bounded by upper and lower limits dictated by comfort, humidity, and equipment safety concerns. However, within these bounds, the chilled-water temperature may not always be low enough to maintain supply air set-point temperatures for the cooling coils. This situation might occur at high loads when the chilled-water flow is at a maximum and is detectable by monitoring the coil discharge air temperatures. Limits on the miimum pressure differential set point might also be imposed to ensure adequate controllability of the cooling-coil control valves.

Implementation

Before commissioning, the parameters of the open-loop control for chilled-water set point [Equation (32)] must be estimated using the results of Table 5. After the system is in operation, these parameters may be fine-tuned with measurements as outlined in the section on Control Optimization Methods. With the parameters specified, the control algorithm is separated into two reset strategies: chilled-water temperature and pressure differential.

Chilled-Water Temperature Reset. The chilled-water supply temperature set point is reset at fixed decision intervals (e.g., 15 min) using the following procedure:

1. Determine the time-averaged position of (or controller output for) the cooling-coil water valves and corresponding discharge air temperatures for "representative" air handlers over the previous decision interval.
2. If more than one valve is saturated at 100% open and their corresponding discharge air temperatures are greater than set point (e.g., 1°F [0.5 K]), then decrease the chilled-water temperature by a fixed amount (e.g., 0.5°F [0.25 K]) and go to step 5. Otherwise, go to step 3.
3. Determine the total chilled-water flow and load.
4. Estimate an optimal chilled-water set point with Equations (9) and (10). Increase or decrease the actual set point in the direction of the near-optimal value by a fixed amount (e.g., 1°F [0.5 K]).
5. Limit the new set point between upper and lower constraints dictated by comfort and equipment safety.

Pump Sequencing. Secondary pumps should be brought online or offline at fixed decision intervals (e.g., 15 min) with the following logic:

1. Evaluate the time-averaged pump controller output over the previous decision interval.
2. If the pump controller is saturated at 100%, then bring the next pump online. Otherwise, go to step 3.
3. If the pump control output is significantly less (e.g., 5%) than the value associated with the first time interval after the last pump was brought online, then bring that pump offline.

Differential Pressure Reset. The set point for differential pressure between supply and return lines should be reset at smaller time intervals than the supply water temperature reset and pump sequencing strategies (e.g., 5 min) using the following procedure:

1. Check the water valve positions (or controller output) for "representative" air handlers and determine the time-averaged values over the last decision interval.
2. If more than one valve has been saturated at 100% open, then increase the differential pressure set point by a fixed value (e.g., 5% of the design value) and go to step 4. Otherwise, go to step 3.
3. If none of the valves have been saturated, then decrease the differential pressure set point by a fixed value (e.g., 5% of the design value).
4. Limit the differential pressure set point between upper and lower constraints.

SEQUENCING AND LOADING MULTIPLE CHILLERS

Multiple chillers are normally configured in parallel and typically controlled to give identical chilled-water supply temperatures. In most cases, controlling for identical set temperatures is the best and simplest strategy. With this approach, the relative loading on operating chillers is controlled by the relative chilled-water flow rates. Typically, the distribution of flow rates to heat exchangers for both chilled and condenser water are dictated by chiller pressure drop characteristics and may be adjusted through flow balancing, but are not controlled using a feedback controller. In addition to the distribution of chilled and condenser water flow rates, the chiller sequencing affects energy consumption. Chiller sequencing defines the conditions under which chillers are brought online and offline. Simple guidelines may be established for each of these controls to provide near-optimal operation.

Near-Optimal Condenser Water Flow Distribution

In general, the condenser water flow to each chiller should be set to give identical leaving condenser water temperatures. This condition approximately corresponds to relative condenser flow rates equal to the relative loads on the chillers, even if the chillers are loaded unevenly. Figure 33 shows results for four sets of two chillers operated in parallel. The curves represent data from chillers at three different installations: (1) a 5500 ton (19 300 kW) variable-speed chiller at the Dallas-Ft. Worth airport, Texas (Braun et al. 1989b); (2) a 550 ton (1930 kW) fixed-speed chiller at an office building in Atlanta (Hackner et al. 1984, 1985); and (3) a 1250 ton (4400 kW) fixed-speed chiller at a large office building in Charlotte, North Carolina (Lau et al. 1985). The capacities of the chillers in the two office buildings were scaled up for comparison with the Dallas-Ft. Worth airport chiller.
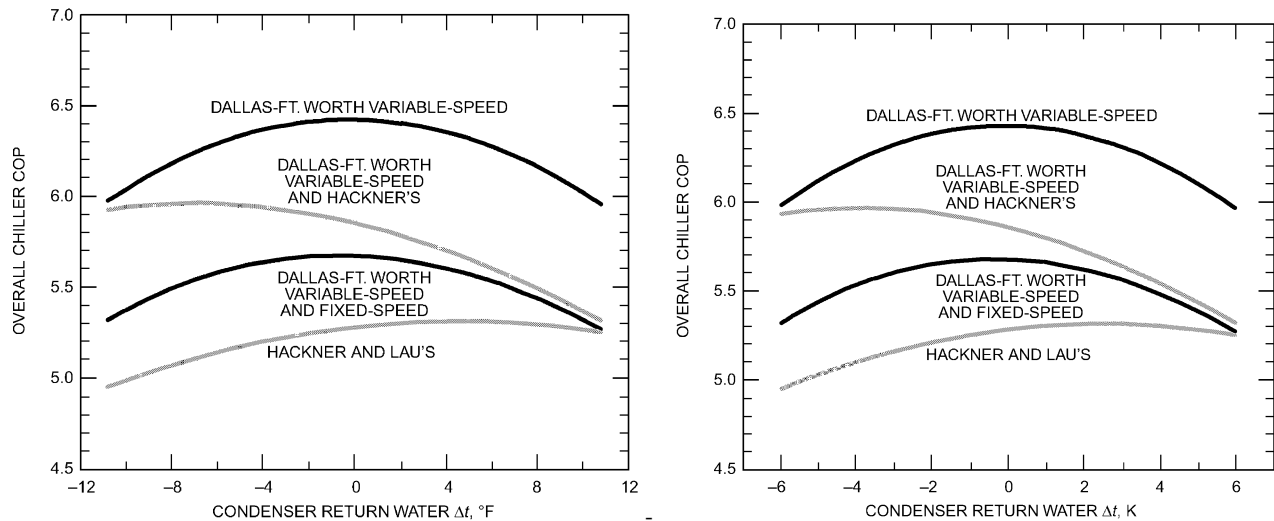
Fig. 33   Effect of Condenser Water Flow Distribution for Two Chillers In Parallel

The overall chiller coefficient of performance (COP) is plotted versus the difference between the condenser water return temperatures for equal chiller loading. For multiple chillers having similar performance characteristics (either variable- or fixed-speed), it is best to distribute the condenser water flow rates so that each chiller has the same leaving condenser water temperature. For situations where chillers do not have identical performance, equal leaving condenser water temperatures result in chiller performance that is close to the optimum. Even for variable- and fixed-speed chiller combinations that have very different performance characteristics, the penalty associated with using identical condenser leaving-water temperatures is small. To achieve equal condenser leaving-water temperatures, it is necessary to properly balance the condenser water flow rates at design operating conditions.

Optimal Chiller Load Distribution

Assuming identical chilled-water return and chiller supply temperatures, the relative chilled-water load for each parallel chiller (load divided by total load) that is operating could be controlled by its relative chilled-water flow rate (flow divided by total flow). To change the relative loadings in response to operating conditions, the individual flow rates must be controlled. However, this is typically not done and it is probably sufficient to establish the load distributions based on design information and then balance the flow rates to achieve these load distributions. Alternatively, the individual chiller loads can be precisely controlled through variation of individual chiller supply water set points.

Chillers with Similar Performance Characteristics. Braun et al. (1989b) showed that, for chillers with identical design COPs and part-load characteristics, a minimum or maximum power consumption occurs when each chiller is loaded according to the ratio of its capacity to the total capacity of all operating chillers. This is equivalent to each chiller operating at equal part-load ratios (load divided by cooling capacity at design conditions). For the ith chiller, the optimal chiller loading is then

$$\dot{Q}_{ch,i}^{*} = \frac{\dot{Q}_{load}}{\displaystyle\sum_{i=1}^{N} \dot{Q}_{ch,des,i}} \, \dot{Q}_{ch,des,i}$$

$$(35)$$

where $\dot{Q}_{load}$ is the total chiller load, $\dot{Q}_{ch,des,i}$ is the cooling capacity of the ith chiller at design conditions, and N is the number of chillers operating.

The loading determined with Equation (35) could result in either minimum or maximum power consumption. However, this solution gives a minimum when the chillers are operating at loads greater than the point at which the maximum COP occurs (i.e., chiller COP decreases with increased loading). Typically, the maximum COP occurs at loads that are less than chiller design capacity.
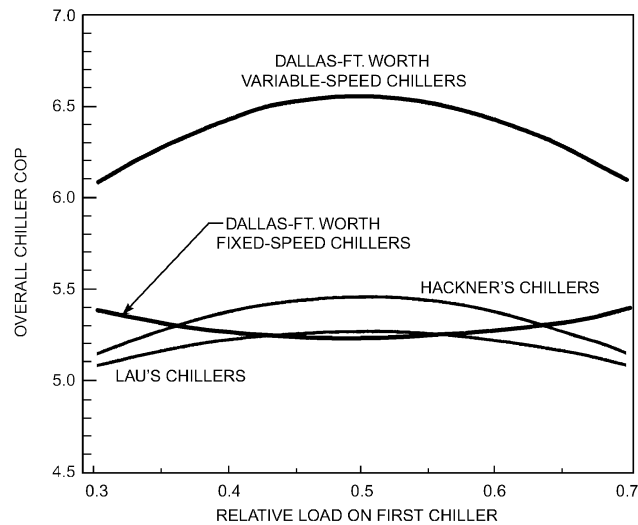
Fig. 34 Effect of Relative Loading for Two Identical Parallel Chillers

Figure 34 shows the effect of relative loading on chiller COP for different sets of identical chillers loaded at approximately 70% of their total capacities. Three of the chiller sets have maximum COPs when evenly loaded [matching the criterion of Equation (35)], whereas the fourth (Dallas-Ft. Worth fixed-speed) obtains a minimum at that point. The part-load characteristic of the Dallas-Ft. Worth fixed-speed chiller is unusual in that the maximum overall COP occurs at its maximum capacity. This chiller was retrofitted with a different refrigerant and drive motor, which derated its capacity from 8700 to 5500 tons (30 6000 to 19 300 kW). As a result, the evaporators and condensers are oversized for its current capacity. Overall, the penalty associated with equally loading the Dallas-Ft. Worth fixed-speed chillers is small compared with optimal loading, and this strategy is probably appropriate. However, a slight reduction in energy consumption is possible if one of the two chillers operates at full capacity. The loading criterion of Equation (35) also works well for many combinations of chillers with different performance characteristics.

To achieve specified relative chiller loadings with equal chilled-water set points, chilled-water flow rates must be properly balanced. The relative loadings of Equation (35) only depend on design information, and flow balancing can be achieved through proper design and commissioning.

Chillers with Different Performance Characteristics. For the general case of chillers with significantly different part-load characteristics, a point of minimum or maximum overall power occurs where the partial derivatives of the individual chiller's power consumption with respect to their loads are equal:

$$\frac{\partial P_{ch,i}}{\partial \dot{Q}_{ch,i}} = \frac{\partial P_{ch,j}}{\partial \dot{Q}_{ch,j}} \qquad \text{for all } i \text{ and } j \qquad (36)$$

and subject to the constraint that

$$\sum_{i=1}^{N} Q_{ch,i} = \dot{Q}_{load} \qquad (37)$$

where $\dot{Q}_{ch,i}$ is the cooling load for the ith chiller and $\dot{Q}_{load}$ is the total cooling load.

In general, the power consumption of a chiller can be correlated as a quadratic function of cooling load and difference between the leaving condenser water and chilled-water supply temperatures according to

$$P_{ch,i} = a_{0,i} + a_{1,i}(t_{cwr,i} - t_{chws,i}) + a_{2,i}(t_{cwr,i} - t_{chws,i})^2$$

$$+ a_{3,i}\dot{Q}_{ch,i} + a_{4,i}\dot{Q}_{ch,i}^2 + a_{5,i}(t_{cwr,i} - t_{chws,i})\dot{Q}_{ch,i} \qquad (38)$$

where, for the ith chiller, t_{cwr} is the leaving condenser water temperature and t_{chws} is the chilled-water supply temperature. The coefficients of Equation (38) ($a_{0,i}$ to $a_{5,i}$) can be determined for each chiller through regression applied to measured or manufacturers' data.

If each chiller has identical leaving condenser and chilled-water supply temperatures, the criterion of Equation (36) applied to the correlation of Equation (38) leads to

$$a_{3,i} + 2a_{4,i}\dot{Q}_{ch,i}^* + a_{5,i}(t_{cwr} - t_{chws})$$

$$= a_{3,j} + 2a_{4,j}\dot{Q}_{ch,j}^* + a_{5,j}(t_{cwr} - t_{chws}) \qquad \text{for } i \neq j \qquad (39)$$

where $\dot{Q}^*_{ch,i}$ is the optimal load for the ith chiller.

Equations (39) and (37) represent a system of N linear equations in terms of N chiller loads that can be solved to give minimum (or possibly maximum) power consumption. For a given combination of chillers, the solution depends on the operating temperatures and total load. However, the individual chiller loads must be constrained to be less than the maximum chiller capacity at these conditions. If an individual chiller load determined from these equations is greater than its cooling capacity, then this chiller should be fully loaded and Equations (39) and (37) should be resolved for the remaining chillers [Equation (39) should only include unconstrained chillers].

To control individual chiller loads with identical chilled-water supply temperatures, individual chilled-water flow rates need to be controlled with two-way valves, which is not typical. However, the distribution of chiller loads could be changed for a fixed-flow distribution by using different chilled-water set-point temperatures. For a given flow and load distribution, the individual chiller set point for parallel chillers is determined according to

$$t_{chws,i} = t_{chwr} - \frac{\dot{Q}_{ch,i}}{f_{F,i}\dot{Q}_{load}}(t_{chwr} - t^*_{chws})$$

(40)

where $f_{F,i}$ is the flow for the ith chiller divided by total flow, t^*_{chws} is the chilled-water supply temperature set point for the combination of chillers determined using the previously defined reset strategies, and t_{chwr} is the temperature of water returned to chillers from the building.

Substituting Equation (40) into Equation (38) and then applying the criterion of Equation (36) leads to the following:

$$A_i + B_i\dot{Q}^*_{ch,i} = A_j + B_j\dot{Q}^*_{ch,j} \qquad \text{for } i \neq j$$

(41)

where

$$A_i = a_{3,i} + [a_{1,i} + 2a_{2,i}(t_{cwr} - t_{chwr})]\frac{t_{chwr} - t^*_{chws}}{f_{F,i}\dot{Q}_{load}}$$

$$+ a_{5,i}(t_{cwr} - t^*_{chws})$$

$$B_i = 2\left[a_{4,i} + a_{5,i}\frac{t_{chwr} - t^*_{chws}}{f_{F,i}\dot{Q}_{load}} + a_{2,i}\left(\frac{t_{chwr} - t^*_{chws}}{f_{F,i}\dot{Q}_{load}}\right)^2\right]$$

Optimal chiller loads are determined by solving the linear system of equations represented by Equations (41) and (37). The individual chiller set points are then evaluated with Equation (40). If any set points are less than the minimum set point or greater than the maximum set point, then the set point should be constrained and Equations (41) and (37) should be resolved for the remaining chillers [Equation (41) should only include unconstrained chillers].

Example 3. Determine the optimal loading for two chillers using the three methods outlined in this section. Table 6 gives the design cooling capacities and coefficients of the curve-fit of Equation (38) for the two chillers. The chillers are operating with a total cooling load of 1440 tons (5064 kW), condenser water return temperature of 85°F (29°C), an overall chilled-water supply temperature set point of 45°F (7°C), and a chilled-water return temperature of 55°F (13°C). Figure 35 shows the COPs for the two chillers as a function load relative to their design loads for the given operating temperatures. Chiller 1 is more efficient at higher part-ratios and less efficient at lower part-load ratios as compared with chiller 2.
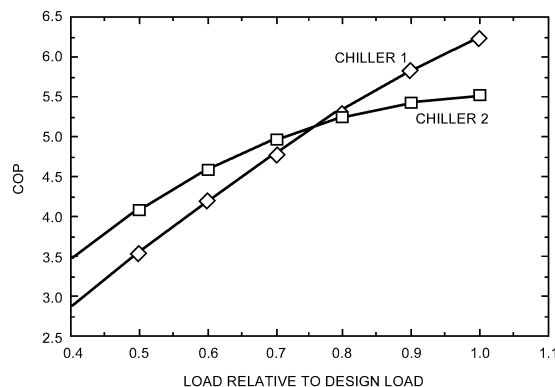

Fig. 35   Chiller COP for Two Chillers

Table 6   Chiller Characteristics for Optimal Loading Example 3

Variable	Units	Chiller 1	Chiller 2

$\dot{Q}_{ch,des,i}$	tons	1250	550
$a_{0,i}$	kW	106.4	119.7
$a_{1,i}$	kW/°F	6.147	0.1875
$a_{2,i}$	kW/°F²	0.1792	0.04789
$a_{3,i}$	kW/ton	−0.0735	−0.3673
$a_{4,i}$	kW/ton²	0.0001324	0.0005324
$a_{5,i}$	kW/ton·°F	−0.001009	0.008526

Table 6   Chiller Characteristics for Optimal Loading Example 3

Variable	Units	Chiller 1	Chiller 2
$\dot{Q}_{ch,des,i}$	kW	4396	1934
$a_{0,i}$	kW	106.4	119.7
$a_{1,i}$	kW/K	11.06	0.3376
$a_{2,i}$	kW/K²	0.5806	0.1552
$a_{3,i}$	kW/kW	−0.0208	−0.1045
$a_{4,i}$	kW/kW²	0.0000107	0.00004304
$a_{5,i}$	kW/(kW·K)	−0.0005163	0.004364

Solution:

First, consider operating the chillers at equal part-load ratios. The ratio of the cooling load to the cooling capacity of the operating chillers is 1440/(1250 + 550) = 0.8 [5064/(4396 + 1934) = 0.8]. From Equation (35), the individual chiller loads are

$$\dot{Q}^{*}_{ch,1} = 0.8(1250 \text{ tons}) = 1000 \text{ tons} \qquad (\dot{Q}^{*}_{ch,1} = 0.8(4396 \text{ kW}) = 3517 \text{ kW})$$

$$\dot{Q}^{*}_{ch,2} = 0.8(550 \text{ tons}) = 440 \text{ tons} \qquad (\dot{Q}^{*}_{ch,2} = 0.8(1934 \text{ kW}) = 1547 \text{ kW})$$

The power for each chiller is computed for the specified operating conditions with Equation (38) and the coefficients of Table 6. For the case of equal part-load ratios, the total chiller power consumption is

$$P_{ch} = P_{ch,1} + P_{ch,2} = 657.5 \text{ kW} + 295.3 \text{ kW} = 952.8 \text{ kW} \qquad (P_{ch} = P_{ch,1} + P_{ch,2} = 649.7 \text{ kW} + 292.2 \text{ kW} = 941.9 \text{ kW})$$

A second solution is determined for optimal chiller loads for the case of equal chilled-water temperature set points and controllable flow for each chiller. In this case, algebraic manipulation of Equations (37) and (39) produces the following results for the individual chiller loads:

$$\dot{Q}^{*}_{ch,1} = \frac{(a_{3,2} - a_{3,1}) + 2a_{4,2}\dot{Q}_{load} + (a_{5,2} - a_{5,1})(t_{cwr} - t_{chws})}{2(a_{4,1} + a_{4,2})} \qquad \dot{Q}^{*}_{ch,1} = \frac{(a_{3,2} - a_{3,1}) + 2a_{4,2}\dot{Q}_{load} + (a_{5,2} - a_{5,1})(t_{cwr} - t_{chws})}{2(a_{4,1} + a_{4,2})}$$

$$= 1219 \text{ tons} \qquad\qquad = 4277 \text{ kW}$$

$$\dot{Q}_{ch,2} = \dot{Q}_{load} - \dot{Q}_{ch,1} = 221 \text{ tons} \qquad (\dot{Q}_{ch,2} = \dot{Q}_{load} - \dot{Q}_{ch,1} = 787 \text{ kW})$$

The resulting power consumption is then

$$P_{ch} = P_{ch,1} + P_{ch,2} = 696.9 \text{ kW} + 224 \text{ kW} = 920.9 \text{ kW} \qquad (P_{ch} = P_{ch,1} + P_{ch,2} = 688.9 \text{ kW} + 222.2 \text{ kW} = 910.8 \text{ kW})$$

Optimal loading of the chillers reduces the overall chiller power consumption by about 4% through heavier loading of chiller 1 and lighter loading of chiller 2 (see Figure 35).

Finally, optimal chiller loading is determined for the case where the individual loadings are controlled by using different chilled-water temperature set points (individual flow is not controllable). To apply Equations (40) and (41), the relative chilled-water flow rate for each chiller must be known. For this example, the relative flow for the ith chiller is assumed to be equal to the ratio of its design capacity to the design capacity for the operating chillers, so that

$$f_{F,1} = \frac{\dot{Q}_{ch,des,1}}{\dot{Q}_{ch,des,1} + \dot{Q}_{ch,des,2}} = \frac{1250}{1250 + 550} = 0.694 \qquad (f_{F,1} = \frac{\dot{Q}_{ch,des,1}}{\dot{Q}_{ch,des,1} + \dot{Q}_{ch,des,2}} = \frac{4396}{4396 + 1934} = 0.694)$$

$$f_{F,2} = \frac{\dot{Q}_{ch,des,2}}{\dot{Q}_{ch,des,1} + \dot{Q}_{ch,des,2}} = \frac{550}{1250 + 550} = 0.306 \qquad (f_{F,2} = \frac{\dot{Q}_{ch,des,2}}{\dot{Q}_{ch,des,1} + \dot{Q}_{ch,des,2}} = \frac{4396}{4396 + 1934} = 0.306)$$

Then, solving Equations (37) and (41) leads to the following results for the individual chiller loads:

$$\dot{Q}_{ch,1}^* = \frac{(A_2 - A_1) + B_2 \dot{Q}_{load}}{B_1 + B_2}$$
$$= 1153 \text{ tons}$$

$$\dot{Q}_{ch,2} = \dot{Q}_{load} - \dot{Q}_{ch,1} = 287 \text{ tons}$$

$$(\dot{Q}_{ch,1}^* = \frac{(A_2 - A_1) + B_2 \dot{Q}_{load}}{B_1 + B_2}$$
$$= 4034 \text{ kW}$$

$$\dot{Q}_{ch,2} = \dot{Q}_{load} - \dot{Q}_{ch,1} = 1030 \text{ kW})$$

These loads lead to a total chiller power consumption of

$$P_{ch} = P_{ch,1} + P_{ch,2} = 713.8 \text{ kW} + 218.1 \text{ kW} = 931.9 \text{ kW} \qquad (P_{ch} = P_{ch,1} + P_{ch,2} = 705.8 \text{ kW} + 216.4 \text{ kW} = 922.2 \text{ kW})$$

Individual chilled-water set points are determined from Equation (40) and are 43.5 and 48.5°F (6.1 and 9°C) for chillers 1 and 2, respectively. Power consumption has increased slightly from the case of identical chiller set points and variable flow.

Note that changing either flows or chilled-water set points complicates overall system control as compared with loading the chillers with fixed part-load ratios, and leads to relatively small savings.

Order for Bringing Chillers Online and Offline

For chillers with similar efficiencies, the order in which chillers are brought online and offline may be dictated by their cooling capacities and the desire to provide even runtimes. However, whenever beneficial and possible, chillers should be brought online in an order that minimizes the incremental increase in energy consumption. At a given condition, the power consumption of any chiller can be evaluated using the correlation given by Equation (38), where the coefficients are determined using manufacturers' data or in-situ measurements. Then, the overall power consumption for all operating chillers is

$$P_{ch} = \sum_{i=1}^{N} P_{ch,i}$$

When additional chiller capacity is required (see next section), the projected power of all valid chiller combinations should be evaluated using Equation (38), with the projected load determined by Equation (35) for chillers with similar performance characteristics, or the solution of Equations (36) subject to the constraint in Equation (37) for chillers with significantly different performance characteristics. Valid chiller combinations involve chillers that are not in alarm or locked out, and with load ratios between a low limit (e.g., 30%) and 100%. The best chiller combination to bring on line should result in the smallest increase (or largest decrease) in overall chiller power consumption as estimated with Equations (42) and (38) with chiller loading determined as outlined in the previous section. For systems with dedicated chilled-water and condenser water pumps and/or cooling towers, the associated power of this equipment should be added to Equation (38) to estimate the load of the chiller plus its auxiliary equipment. It is recommended that these estimates be performed online using in-situ measurements of each chiller's discharge chilled-water temperature and entering condenser water temperature, and the projected load of the chiller. A chiller should be shut down when its load drops below the spare capacity load of the current number of online chillers; for a primary/secondary chilled-water system, the primary chilled-water flow will remain above the secondary chilled-water flow once the chiller is shut down. (The spare capacity load is equal to the rated capacity of the online chillers minus the actual measured load of the online chillers.)

For chillers with similar design cooling capacities, a simpler (although suboptimal) approach can be used for determining the order for bringing chillers online and offline. In this case, the chiller with the highest peak COP can be brought online first, followed by the second most efficient chiller, etc., and then brought offline in reverse order. The maximum COP for each chiller can be evaluated using manufacturers' design and part-load data or from curve-fits to in-situ performance.
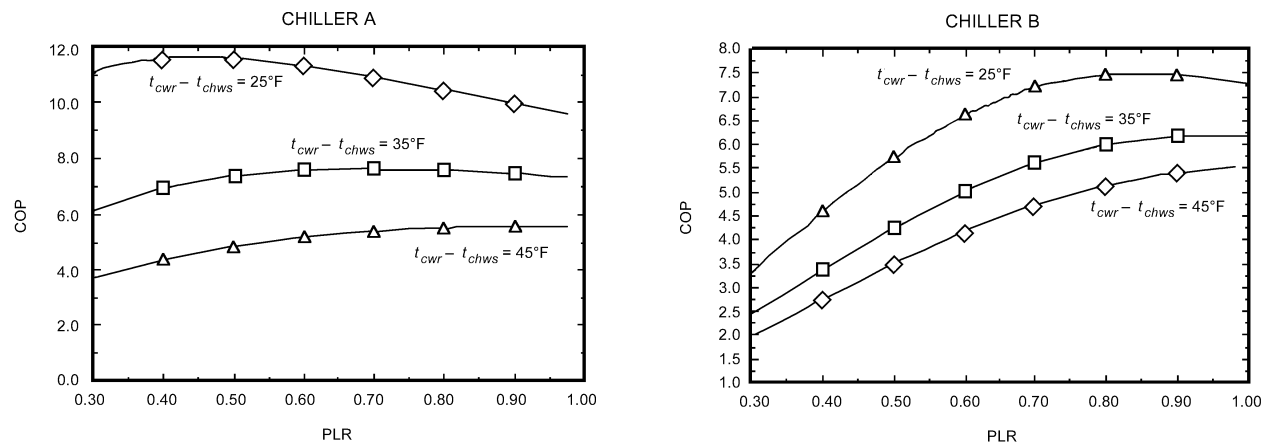

Fig. 36   Chiller A and B Performance Characteristics for Maximum COP, Example 4

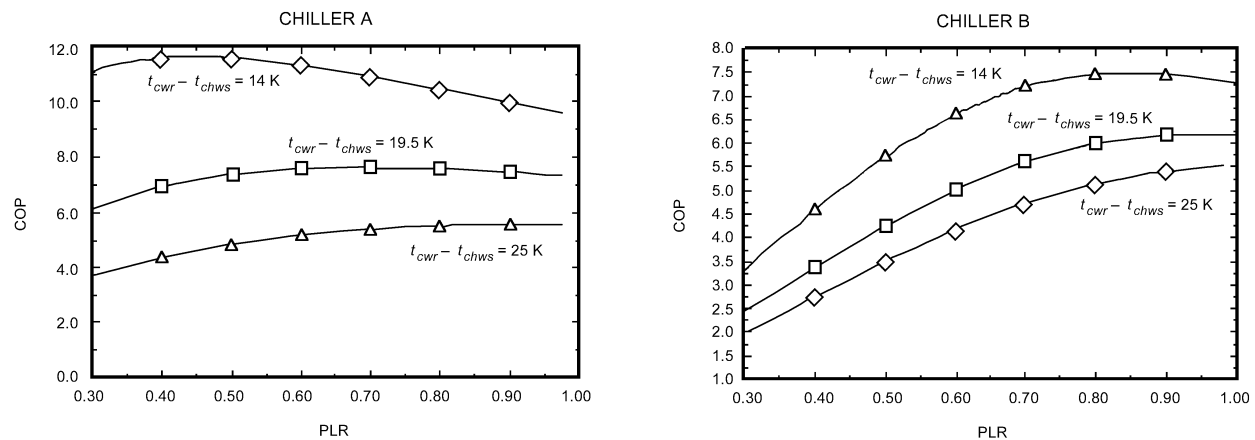DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*

Fig. 36 Chiller A and B Performance Characteristics for Maximum COP, Example 4

The chiller load associated with maximum COP for each chiller can be determined by applying a first-order condition for a maximum, using Equation (38) and the definition of COP. For this functional form, the maximum (or possibly minimum) COP occurs for

$$\dot{Q}^*_{ch,i} = \sqrt{\frac{a_{0,i} + a_{1,i}(t_{cwr} - t_{chws}) + a_{2,i}(t_{cwr} - t_{chws})^2}{a_{4,i}}}$$

The load determined from Equation (43) yields a maximum COP whenever it is real and bounded between upper and lower limits. Otherwise, it can be assumed that the maximum COP occurs at full load conditions. Typically, the maximum COP for centrifugal chillers occurs between about 40 and 80% and, for small multi-compressor or inverter drive chillers, between 15 and 40% of design load. COP increases as the temperature difference between the condenser leaving water and chilled-water supply decreases. Equation (43) could be applied online to determine the rank ordering of chillers to bring online as a function of operating temperatures. However, it is often sufficient to use Equation (43) at the design temperature difference and establish a chiller sequencing order at the design or commissioning stage.

Example 4. Determine the loads for maximum COP for two different chillers at a chilled-water set point of 45°F (7°C) and a condenser water return temperature of 80°F (27°C). Table 7 gives the design cooling capacities and coefficients of the curve-fit of Equation (38) for the two chillers. Figure 36 shows the COPs of the two chillers determined from the correlations as a function of relative load (PLR) and temperature difference ($t_{cwr} - t_{chws}$). These chillers have identical performance at design conditions, but very different part-load characteristics because of different methods used for capacity control.

Solution:

Loading associated with the maximum COP for each chiller is determined using Equation (43) and the coefficients of Table 7. Power for each chiller is then determined using Equation (38) and the COP follows directly. Results of the calculations are given in Table 8. The maximum COP for chiller A is about 20% greater than that for chiller B at the specified operating temperatures and should be brought online first.

Load Conditions for Bringing Chillers Online or Offline

In general, chillers should be brought online at conditions where the total power (including pumps and tower or condenser fans) of operating with the additional chiller would be less than without it. Conversely, a chiller should be taken offline when the total power of operating with that chiller would be less than with it. In practice, the switch point for bringing a chiller online should be greater than that for bringing that same chiller offline (e.g., 10%), to ensure stable control. The optimal sequencing of chillers depends primarily on their part-load characteristics and the manner in which the chiller pumps are controlled.

Table 7 Chiller Characteristics for Maximum COP, Example 4

Variable	Unit	Chiller A	Chiller B
$\dot{Q}_{ch,des,i}$	tons	5421	5421
$a_{0,i}$	kW	262.6	187.2
$a_{1,i}$	kW/°F	−25.36	96.19
$a_{2,i}$	kW/°F²	0.9718	−0.4314
$a_{3,i}$	kW/ton	−0.02568	−0.4314
$a_{4,i}$	kW/ton²	0.00004046	0.0001106
$a_{5,i}$	kW/ton·°F	0.005289	−0.004537

Table 7 Chiller Characteristics for Maximum COP, Example 4

Variable	Unit	Chiller A	Chiller B
$\dot{Q}_{ch,des,i}$	MW	19.07	19.07
$a_{0,i}$	kW	262.6	187.2
$a_{1,i}$	kW/K	−45.65	173.1
$a_{2,i}$	kW/K^2	3.149	−1.398
$a_{3,i}$	kW/kW	−0.007 301	−0.1227
$a_{4,i}$	kW/kW^2	0.000 003 271	0.000 008 944
$a_{5,i}$	kW/(kW·K)	0.002 707	−0.002 322

Table 8 Results for Maximum COP, Example 4

Variable	Chiller A	Chiller B
$\dot{Q}_{ch,i}^{*}$	3738 ton	5229 ton
PLR_i	0.690	0.965
$P_{ch,i}$	1727 kW	2964 kW
COP_i	7.61	6.21

Table 8 Results for Maximum COP, Example 4

Variable	Chiller A	Chiller B
$\dot{Q}_{ch,i}^{*}$	13.65 MW	18.59 MW
PLR_i	0.716	0.975
$P_{ch,i}$	1857 kW	3038 kW
COP_i	7.35	6.12

Dedicated Pumps. Where individual condenser and chilled-water pumps are dedicated to the chiller, Braun et al. (1989b) and Hackner et al. (1985) showed that a chiller should be brought online when the operating chillers reach their capacity. This conclusion is the result of considering both the chiller and pumping power in determining optimal control. If pumping power is ignored, the optimal chiller sequencing occurs when chiller efficiency is maximized at each load. Because maximum efficiency often occurs at part-load conditions, the optimal point for adding or removing chillers may occur when chillers are operating at less than their capacity. However, the additional pumping power required for bringing additional pumps online with the chiller usually offsets any reductions in overall chiller power consumption associated with part-load operation.

When pumps are dedicated to chillers, situations may arise where chillers are operating at less than their capacity but chilled-water flow to the cooling coils is insufficient to meet the building load. This generally results from inadequate design or improper maintenance. Under these circumstances, either some zone conditions need to float to reduce the chilled-water set point (if possible), or an additional chiller needs to be brought online. Monitoring the zone air-handler conditions is one way to detect this situation. If (1) the chilled-water set point is at its lower limit, (2) any air-handler water control valves are saturated at 100% open, and (3) their corresponding discharge air temperatures are significantly greater (e.g., 2°F [1 K]) than set point, then the chilled-water flow is probably insufficient and an additional chiller/pump combination could be brought online. One advantage of this approach is that it is consistent with the reset strategies for both fixed- and variable-speed chilled-water systems.

Chillers can be brought online or offline with the following logic:

1. Evaluate the time-averaged values of the chilled-water supply temperature and overall cooling load over a fixed time interval (e.g., 5 min).
2. If the chilled-water supply temperature is significantly greater than the set point (e.g., 1°F [0.5 K]), then bring the next chiller online. Otherwise, go to step 3.
3. Determine the time-averaged position of the cooling-coil water valves and corresponding discharge air temperatures for "representative" air handlers.
4. If (a) the chilled-water supply set point is at its lower limit, (b) more than one valve is saturated at 100% open, and (c) their corresponding discharge air temperatures are significantly greater than set point (e.g., 1°F [0.5 K]), then bring another chiller/pump combination online. Otherwise, go to step 5.
5. If the cooling load is significantly less (e.g., 10%) than the value associated with the first time interval after the last chiller was brought online, then take that chiller offline.

Nondedicated Pumps. For systems without dedicated chiller pumps (e.g., variable-speed primary systems), the optimal load conditions for bringing chillers online or offline do not generally occur at full chiller capacity. In determining optimal chiller switch

points, ideally both chiller and pumping power should be considered because pressure drop characteristics and pumping change when a chiller is brought online or offline. However, approximate values of optimal switch points may be estimated by considering only chiller power.

A chiller should be brought online whenever it would reduce the overall chiller power or if the current chillers can no longer meet the load (see previous section). A chiller should be added if the power consumption associated with $(N + 1)$ chillers is significantly less (e.g., 5%) than the current N chillers, with both conditions evaluated using Equation (42) with correlations of the form given in Equation (38), and sequencing and loading determined as outlined in previous sections. Conversely, a chiller should be removed if the power consumption associated with the $(N - 1)$ chillers is significantly less (e.g., 5%) than the current N chillers. The decision to add or remove chillers is readily determined using the current load and operating temperatures.

SIMPLIFIED STATIC OPTIMIZATION OF COOLING PLANTS

Optimal supervisory control of cooling equipment may involve determining the control that minimizes the total operating cost. For an all-electric system without significant storage, cost optimization leads to minimization of power at each instant in time. Optimal control depends on time, albeit indirectly through changing cooling requirements and ambient conditions. In this section, the static optimization methods described in the Methods section are simplied and applied to cooling

Simplified System-Based Optimization Approach

The component-based optimization method presented by Braun et al. (1989a) described in the Methods section was used to develop a simpler method for determining optimal control. The method involves correlating overall cooling plant power consumption using a quadratic functional form. Minimizing this function leads to linear control laws for control variables in terms of uncontrolled variables. The technique may be used to tune parameters of the cooling tower and chilled-water reset strategies presented in the section on Supervisory Control Strategies and Tools. It may also be used to define strategies for supply air tem-perature reset for VAV systems and flow control for variable-speed condenser water pumps.

In the vicinity of any optimal control point, plant power consumption may be approximated as a quadratic function of the continuous control variables for each of the operating modes (i.e., discrete control mode). A quadratic function also correlates power consumption in terms of uncontrolled variables (i.e., load, ambient temperature) over a wide range of conditions. This leads to the following general functional relationship between overall cooling plant power and the controlled and uncontrolled variables:

$$J(\mathbf{f}, \mathbf{M}, \mathbf{u}) = \mathbf{u}^T \mathbf{A} \mathbf{u} + \mathbf{b}^T \mathbf{u} + \mathbf{f}^T \mathbf{C} \mathbf{f} + \mathbf{d}^T \mathbf{f} + \mathbf{f}^T \mathbf{E} \mathbf{u} + g \qquad (4)$$

where J is the total plant power, $\mathbf{u}$ is a vector of continuous and free control variables, $\mathbf{f}$ is a vector of uncontrolled variables, $\mathbf{M}$ is a vector of discrete control variables, and the superscript T designates the transpose vector. $\mathbf{A}$, $\mathbf{C}$, and $\mathbf{E}$ are coefficient matrices, $\mathbf{b}$ and $\mathbf{d}$ are coefficient vectors, and g is a scalar. The empirical coefficients of this function depend on the operating modes so that these constants must be determined for each feasible combination of discrete control modes, $\mathbf{M}$.

A solution for the optimal control vector that minimizes power may be determined analytically by applying the first-order condition for a minimum. Equating the Jacobian of Equation (4) with respect to the control vector to zero and solving for the optimal control set points gives

$$\mathbf{u}^* = \mathbf{k} + \mathbf{K} \mathbf{f} \qquad (5)$$

where

$$\mathbf{k} = -\mathbf{A}^{-1} \mathbf{b}/2 \qquad (6)$$

$$\mathbf{K} = -\mathbf{A}^{-1} \mathbf{E}/2 \qquad (7)$$

The cost associated with the unconstrained control of Equation (5) is

$$J^* = \mathbf{f}^T \theta \mathbf{f} + \sigma^T \mathbf{f} + \tau \qquad (8)$$

where

$$\theta = \mathbf{K}^T \mathbf{A} \mathbf{k} + \mathbf{E} \mathbf{K} + \mathbf{C} \qquad (9)$$

$$\sigma = 2\mathbf{K} \mathbf{A} \mathbf{k} + \mathbf{K} \mathbf{b} + \mathbf{E} \mathbf{k} + \mathbf{d} \qquad (10)$$

$$\tau = \mathbf{K}^T \mathbf{A} \mathbf{k} + \mathbf{b}^T \mathbf{k} + g \qquad (11)$$

The control defined by Equation (5) results in a minimum power consumption if $\mathbf{A}$ is positive definite. If this condition holds and if the system power consumption is adequately correlated with Equation (4), then Equation (5) dictates that the optimal continuous, free control variables vary as a nearly linear function of the uncontrolled variables. However, a different linear relationship, i.e. different $\mathbf{A}$, $\mathbf{C}$, $\mathbf{E}$, $\mathbf{b}$, $\mathbf{d}$ and g, applies to each feasible combination of discrete control modes defined by $\mathbf{M}$. The minimum cost associated with each mode combination must be computed from Equation (8) and compared to identify the global minimum.

Uncontrolled Variables. As discussed in the background section, optimal control variables primarily depend on ambient wet-bulb temperature (or dry-bulb temperature in the case of air-cooled chillers) and total chilled-water load. The load affects the heat transfer requirements for all heat exchangers, whereas the wet-bulb temperature affects chilled- and condenser water temperatures necessary to achieve a given heat transfer rate. As discussed in the section on Supervisory Control Strategies and Tools, cooling coil heat transfer depends on the coil entering wet-bulb temperature. However, this reduces to an ambient wet-bulb temperature dependence for a given

ventilation mode (e.g., minimum outside air or economizer) and fixed zone conditions. Thus, separate cost functions are necessary for each ventilation mode, with load and ambient wet bulb as uncontrolled variables. Alternatively, for a specified ventilation strategy (e.g., the economizer strategy from the section on Supervisory Control Strategies and Tools), three uncontrolled variables could be used for all ventilation modes: load, ambient wet-bulb temperature, and average cooling-coil inlet wet-bulb temperature.

For radiant cooling systems, if latent cooling is handled by package dehumidifying equipment (e.g., a dedicated outdoor air system), the chilled-water set point can be raised well above what is needed for dehumidification and the main chiller will operate more efficiently. Gayeski et al. (2011a) and Katipamula et al. (2010b) show that part-load efficiency may benefit greatly from using optimal chilled-water flow rate and temperature in response to part-load ratio.

Additional uncontrolled variables that could be important if varied over a wide range are the individual-zone latent-to-sensible load ratios and the ratios of individual sensible zone loads to the total sensible loads for all zones. However, these variables are difficult to determine from measurements and are of secondary importance.

Free Control Variables. The number of independent or "free" control variables in the optimization can be reduced significantly by using the simplified strategies presented in the section on Supervisory Control Strategies and Tools. For instance, the optimal static pressure set point for a VAV system should keep at least one VAV box fully open and should not be considered as a free optimization variable. Similarly, supply air temperature for a constant-air-volume (CAV) system should be set to minimize reheat. Additional near-optimal guidelines were presented for sequencing of cooling tower fans, sequencing of chillers, loading of chillers, reset of pressure differential set point for variable-speed pumping, and chilled-water reset with fixed-speed pumping. Furthermore, Braun et al. (1989b) showed that using identical supply air set points for multiple air handlers gives near-optimal results for VAV systems.

For all variable-speed auxiliary equipment (i.e., pumps and fans), the free set-point variables to use in Equation (1) could be reduced to the following: (1) supply air set temperature, (2) chilled-water set temperature, (3) tower airflow relative to design capacity, and (4) condenser water flow relative to design capacity. All other continuous supervisory control variables are dependent on these variables with the simplified strategies presented in the section on Supervisory Control Strategies and Tools.

Some of the dependent control variables may be discrete control variables. For instance, variable-flow pumping may be implemented with multiple fixed- and variable-speed pumps, where the number of operating pumps is a discrete variable that changes when a variable-speed pump reaches its capacity. These discrete changes could lead to discrete changes in cost because of changes in overall pump efficiency. However, this has a relatively small effect on overall power consumption and may be neglected in fitting the overall cost function to changes in the control variables.

Some discrete control variables may also be independent variables. In general, different cost functions arise for all operating modes consisting of each possible combination of discrete control variables. With all variable-speed pumps and fans, the only significant discrete control variable is the number of operating chillers. Then, optimization involves determining optimal values of only four continuous control variables for each of the feasible chiller modes. A chiller mode defines which of the available chillers are to be online. The chiller mode giving the minimum overall power consumption represents the optimum. For a chiller mode to be feasible, the specified chillers must operate safely within their capacity and surge limits. In practice, abrupt changes in the chiller modes should also be avoided. Large chillers should not be cycled on or off except when the savings associated with the change are significant.

Using fixed-speed equipment reduces the number of free continuous control variables. For instance, supply air temperature is removed as a control variable for CAV systems, and chilled-water temperature is not included for fixed-speed chilled-water pumping. However, for multiple chilled-water pumps not dedicated to chillers, the number of operating pumps can become a free discrete control variable. Similarly, for multiple fixed-speed cooling tower fans and condenser water pumps, each of the discrete combinations can be considered as a separate mode. However, for multiple cooling tower cells with multiple fan speeds, the number of possible combinations may be large. A simpler approach that works satisfactorily is to treat relative flows as continuous control variables during the optimization and to select the discrete relative flow that is closest to the optimal value. At least three relative flows (discrete flow modes) are necessary for each chiller mode to fit the quadratic cost function. The number of possible sequencing modes for fixed-speed pumps is generally much more limited than that for cooling tower fans, with two or three possibilities (at most) for each chiller mode. In fact, with many current designs, individual pumps are physically coupled with chillers, and it is impossible to operate more or fewer pumps than the number of operating chillers. Thus, it is generally best to treat the control of fixed-speed condenser water pumps with a set of discrete control possibilities rather than use a continuous control approximation.

Training. The coefficients of Equation (4) must be determined empirically, and a variety of approaches have been proposed. One approach is to apply regression techniques directly to measurements of total power consumption. Because the cost function is linear with respect to the empirical coefficients, linear regression techniques may be used. A set of experiments can be performed over the expected range of operating conditions. Large amounts of data that include the entire range must be taken to account for measurement uncertainty. The regression could possibly be performed online using least-squares recursive parameter updating (Ljung and Söderström 1983). However, precautions should be taken to ensure that the matrix **A** is positive definite, which guarantees a minimum. If system power is relatively "flat," automated methods could generate coefficients that produce a maximum in power consumption rather than a minimum (Brandemuehl and Bradford 1998).

Rather than fitting empirical coefficients of the system-cost function of Equation (4), the coefficients of the optimal control Equation (5) and the minimum-cost function of Equation (8) may be estimated directly. At a limited set of conditions, optimal values of the continuous control and free variables may be estimated through trial-and-error variations. Only three independent conditions are necessary to determine coefficients of the linear control law given by Equation (5) if the load and wet bulb are the only uncontrolled variables. The coefficients of the minimum cost function can then be determined from system measurements with the linear control law in effect. The disadvantage of this approach is that there is no direct way to handle physical constraints on the controls.

Summary and Constraint Implementation. The methodology for determining the near-optimal control of a chilled-water system may be summarized as follows:

1. Change the chiller operating mode if system operation is at the limits of chiller operation (near surge or maximum capacity).

2. For the current set of conditions (load and wet bulb), estimate the feasible modes of operation, **M**, that avoid operating the chiller and condenser pump at their limits.
3. For the current operating mode, determine optimal values of the continuous controls using Equation (5).
4. Determine a constrained optimum if controls exceed their bounds.
5. Repeat steps 3 and 4 for each feasible operating mode.
6. Change the operating mode if the optimal cost associated with the new mode is significantly less than that associated with the current mode.
7. Change the values of the continuous control variables. When treating multiple-speed fan control with a continuous variable, use the discrete control closest to the optimal continuous value.

If the linear optimal control Equation (5) is directly determined from optimal control results, then the constraints on controls may be handled directly. Otherwise, a simple solution is to constrain the individual control variables as necessary and neglect the effects of the constraints on the optimal values of the other controls and the minimum cost function. The variables of primary concern for constraints are the chilled-water and supply air set temperatures. These controls must be bounded for proper comfort and safe operation of the equipment. On the other hand, cooling tower fans and condenser water pumps should be sized so the system performs efficiently at design loads, and constraints on control of this equipment should only occur under extreme conditions.

The optimal value of the chilled-water supply temperature is coupled to the optimal value of the supply air temperature, so decoupling these variables in evaluating constraints is generally not justified. However, optimization studies indicate that when either control is operated at a bound, the optimal value of the other free control is approximately bounded at a value that depends only on the ambient wet-bulb temperature. The optimal value of this free control (either chilled-water or supply air set point) may be estimated at the load at which the other control reaches its limit. Coupling between optimal values of the chilled-water and condenser water loop controls is not as strong; interactions between constraints on these variables may be neglected.

Case Studies. Braun et al. (1987) correlated the power consumption of the Dallas-Ft. Worth airport chiller, condenser pumps, and cooling tower fans with the quadratic cost function given by Equation (4) and showed good agreement with data. Because the chilled-water loop control was not considered, the chilled-water set point was treated as a known uncontrolled variable. The discrete control variables associated with the four tower cells with two-speed fans and the three condenser pumps were treated as continuous control variables. The optimal control determined by the near-optimal Equation (5) also agreed well with that determined using a nonlinear optimization applied to a detailed simulation of the system.

In subsequent work, Braun et al. (1989a) considered complete system simulations (cooling plant and air handlers) to evaluate the performance of the quadratic, system-based approach. A number of different system characteristics were considered. Figures 4, 5, 22, and 23 show comparative results between the controls as determined with the component-based and system-based methods for a range of loads, for a relatively low and high ambient wet-bulb temperature (60°F and 80°F [16°C and 27°C]).
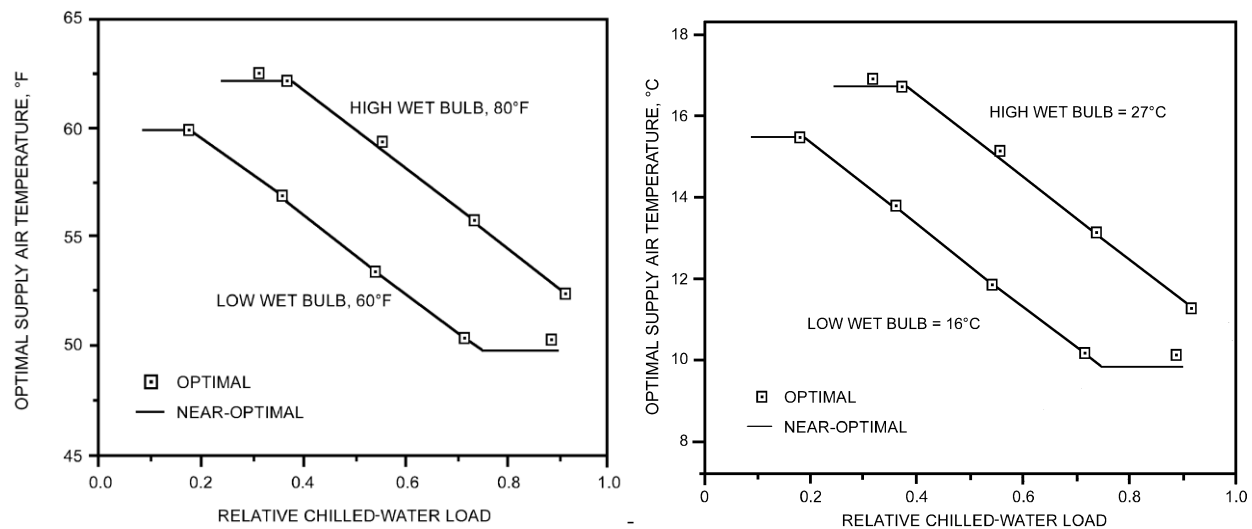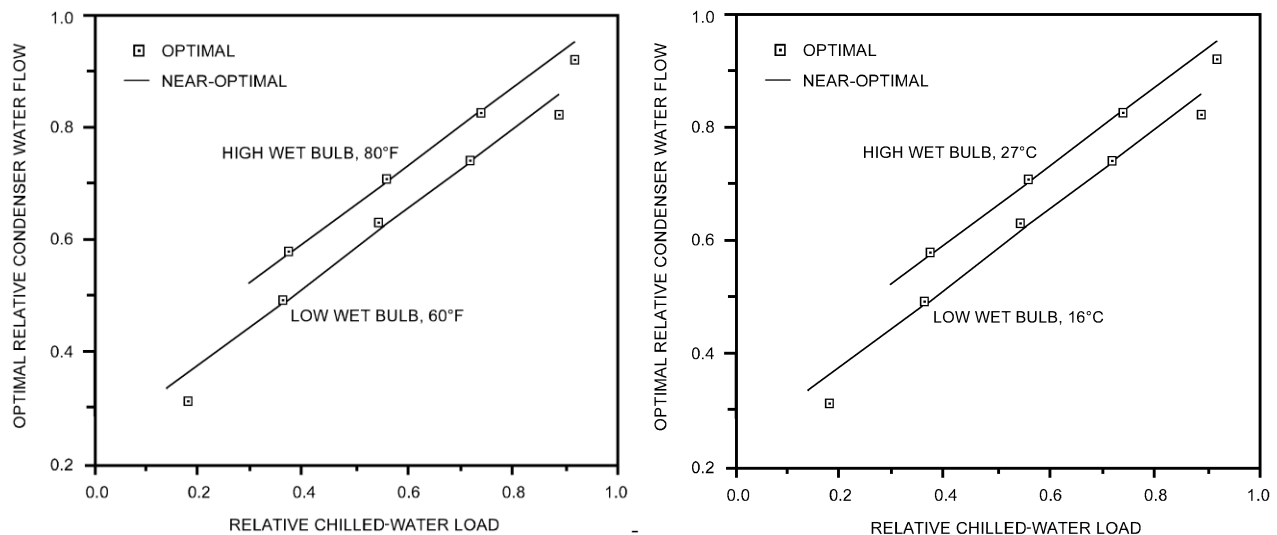

Fig. 4   Comparisons of Optimal Supply Air Temperature

Fig. 5   Comparisons of Optimal Condenser Pump Control

In Figures 4 and 22, optimal values of the chilled-water and supply air temperatures are compared for a system with variable air and water flow. The near-optimal control equation provides a good fit to the optimization results for all conditions considered. The chilled-water temperature was constrained between 38 and 55°F (3 and 13°C), while the supply air set point was allowed to float freely. Figures 4 and 22 show that, for the conditions where the chilled-water temperature is constrained, the optimal supply air temperature is also nearly bounded at a value that depends on the ambient wet bulb.

Optimal relative cooling tower air and condenser water flow rates are compared in Figures 5 and 23 for a system with variable-speed cooling tower fans and condenser water pumps. Although the optimal controls are not exactly linear functions of the load, the linear control equation provides an adequate fit. The differences in these controls result in insignificant differences in overall power consumption, because, as discussed in the background section, the optimum is extremely flat with respect to these variables. The nonlinearity of the condenser loop controls is partly due to the constraints imposed on the chilled-water set temperature. However, this effect is not very significant. Figures 5 and 23 also suggest that the optimal condenser loop control is not very sensitive to ambient wet-bulb temperature.

Static Optimization for Cooling Plants

The cooling system illustrated in Figure 1 of the Methods section depicts multiple chillers, cooling towers, and pumps providing chilled water to air-handling units to cool air that is supplied to building zones. Although cooling needs at any given time may be met with different modes of operation and set points, only one set of controls and mode results in minimum power consumption. This optimal control point results from tradeoffs between the energy consumption of different components. For instance, increasing the number of cooling tower cells (or increasing fan speeds) increases fan power but reduces chiller power because the temperature of the water supplied to the chiller's condenser is decreased. Similarly, increasing condenser water flow by adding pumps (or increasing pump speed) decreases chiller power but increases pump power.

Similar tradeoffs exist for the chilled-water loop variables of systems with variable-speed chilled-water pumps and air handler fans. For instance, increasing the chilled-water set point reduces chiller power but increases pump power because greater flow is needed to meet the load. Increasing the supply air set point increases fan power, but decreases pump power.

Figure 8 illustrates the sensitivity of the total power consumption to condenser water-loop controls (from Braun et al. 1989b) for a single chiller load, ambient wet-bulb temperature, and chilled-water supply temperature. Contours of constant power consumption are plotted versus cooling tower fan and condenser water pump speed for a system with variable-speed fans and pumps. Near the optimum, power consumption is not sensitive to either of these control variables, but increases significantly away from the optimum. The rate of increase in power consumption is particularly large at low condenser pump speeds. A minimum pump speed is necessary to overcome the static pressure associated with the height of the water discharge in the cooling tower above the sump. As the pump speed approaches this value, condenser flow approaches zero and chiller power increases dramatically. A pump speed that is too high is generally better than one that is too low. The broad area near the optimum indicates that, for a given load, the optimal setting does not need to be accurately determined. However, optimal settings change significantly when there are widely varying chiller loads and ambient wet-bulb temperature.
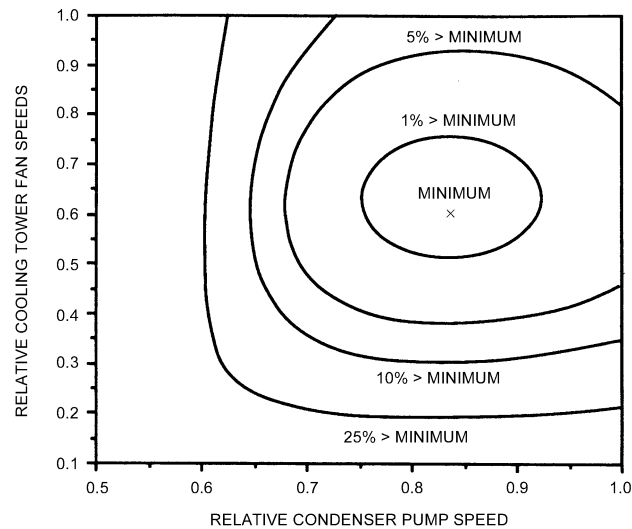
Fig. 8 Example Chiller Plant Power Contours for Condenser-Loop Control Variables

Figure 9 illustrates the sensitivity of power consumption to chilled-water and supply air set-point temperatures for a system with variable-speed chilled-water pumps and air handler fans (Braun et al. 1989b). Within about 3°F (2 K) of the optimum values, power consumption is within 1% of the minimum. Outside this range, sensitivity to the set points increases significantly. The penalty associated with operation away from the optimum is greater in the direction of smaller differences between the supply air and chilled-water set points. As this temperature difference is reduced, the required flow of chilled water to this coil increases and the chilled-water pumping power is greater. For a given chilled-water or supply air temperature, the temperature difference is limited by the heat transfer characteristics of the coil. As this limit is approached, the required water flow and pumping power would become infinite if the pump speed were not constrained. It is generally better to have too large rather than too small a temperature difference between the supply air and chilled-water set points.
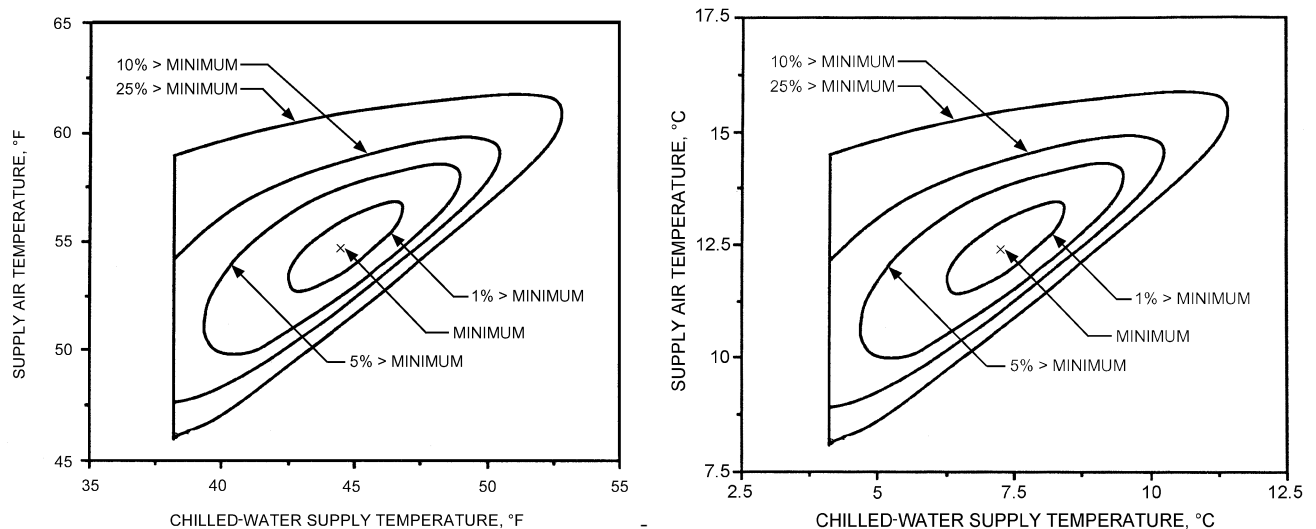
Fig. 9 Example Chiller Plant Power Contours for Chilled-Water and Supply Air Temperatures

For constant chilled-water flow, tradeoffs in energy use with chilled-water set point are very different than for variable-flow systems. Increasing the chilled-water set point reduces chiller power consumption, but has little effect on chilled-water pumping energy. Therefore, the benefits of chilled-water temperature reset are more significant than for variable-flow systems (although variable-flow systems use less energy). For constant chilled-water flow, the minimum-cost strategy is to raise the chilled-water set point to the highest value that will keep all discharge air temperatures at their set points and keep zone humidities within acceptable bounds.

For constant-volume (CAV) air-handling systems, tradeoffs in energy use with supply air set point are also very different than for variable-air-volume systems. Increasing the supply air set point for cooling reduces both the cooling load and reheat required, but does

not change fan energy. Again, the benefits of supply air temperature reset for CAV systems are more significant than for VAV systems (although VAV systems use less energy). In general, the set point for a CAV system should be at the highest value that will keep all zone temperatures at their set points and all humidities within acceptable limits.

In addition to the set points used by local-loop controllers, a number of operational modes can affect performance. For instance, significant energy savings are possible when a system is properly switched over to an economizer cycle. At the onset of economizer operation, return dampers are closed, outdoor air dampers are opened, and the maximum possible outdoor air is supplied to cooling coils. Two different types of switchover are typically used: (1) dry-bulb and (2) enthalpy. With a dry-bulb economizer, the switchover occurs when the ambient dry-bulb temperature is less than a specified value, typically between 55 and 65°F (13 and 18°C). With an enthalpy economizer, the switchover typically happens when the outdoor enthalpy (or wet-bulb temperature) is less than the enthalpy (or wet-bulb temperature) of the return air. Although the enthalpy economizer yields lower overall energy consumption, it requires wet-bulb temperature or dry-bulb and relative humidity measurements.

Another important operation mode is the sequencing of chillers and pumps. Sequencing defines the order and conditions associated with bringing equipment online or offline. Optimal sequencing depends on the individual design and part-load performance characteristics of the equipment. For instance, more-efficient chillers should generally be brought online before less-efficient ones. Furthermore, the conditions where chillers and pumps should be brought online depend on their performance characteristics at part-load conditions.

Figure 10 shows an example of optimal system performance (i.e., optimal set-point choices) for different combinations of chillers and fixed-speed pumps in parallel as a function of load relative to the design load for a given ambient wet-bulb temperature. For this system (from Braun et al. 1989b), each component (chillers, chilled-water pumps, and condenser water pumps) in each parallel set is identical and sized to meet half of the design requirements. The best performance occurs at about 25% of the design load with one chiller and pump operating. As load increases, system COP decreases because of decreasing chiller COP and a nonlinear increase in the power consumption of cooling tower and air handler fans. A second chiller should be brought online at the point where the overall COP of the system is the same with or without the chiller. For this system, this optimal switch point occurs at about 38% of the total design load or about 75% of the individual chiller's capacity. The optimal switch point for bringing a second condenser and chilled-water pump online occurs at a much higher relative chilled load (0.62) than the switch point for adding or removing a chiller (0.38). However, pumps are typically sequenced with chillers (i.e., they are brought online together). In this case, Figure 10 shows that the optimal switch point for bringing a second chiller online (with pumps) is about 50% of the overall design load or at the design capacity of the individual chiller. This is generally the case for sequencing chillers with dedicated pumps.
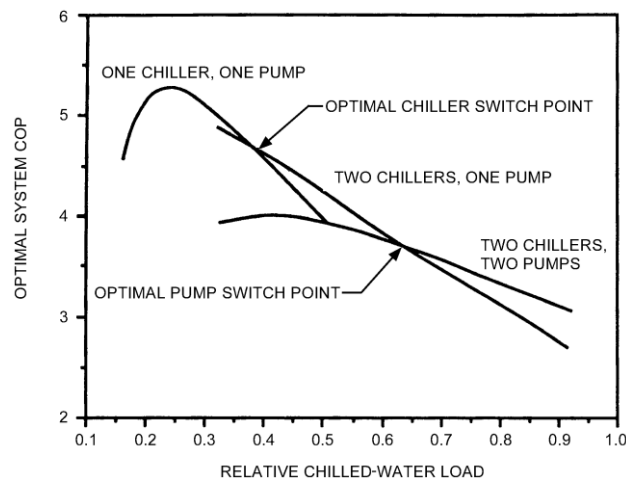
Fig. 10   Example of Effect of Chiller and Pump Sequencing on Optimal Performance

In most cases, zone humidities are allowed to float between upper and lower limits dictated by comfort (see Chapter 8 of the 2009 *ASHRAE Handbook—Fundamentals*). However, VAV systems can control zone humidity and temperature simultaneously. For a zone being cooled, the equipment operating costs are minimized when the zone temperature is at the upper bound of the comfort region. However, operating simultaneously at the upper limit of humidity does not minimize operating costs. Figure 11 shows an example comparison of system COP and zone humidity associated with fixed and free-floating zone humidity as a function of the relative load (from Braun et al. 1989b). Over the range of loads, allowing the humidity to float within the comfort zone produces a lower cost and zone humidity than setting the humidity at the highest acceptable value. The largest differences occur at the highest loads. Operation with the zone at the upper humidity bound results in lower latent loads than with a free-floating humidity, but this humidity control constraint requires a higher supply air temperature, which in turn results in greater air handler power consumption. For minimum energy costs, the humidity should be allowed to float freely within the bounds of human comfort.
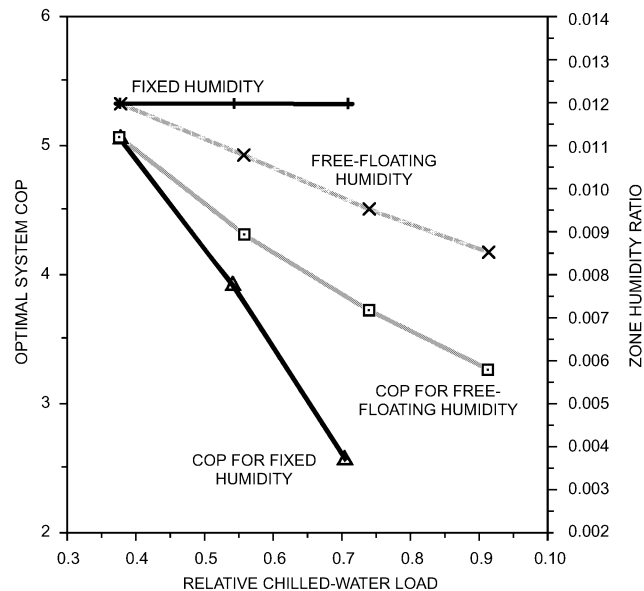
Fig. 11   Example Comparison of Free-Floating and Fixed Humidity

Effects of Load and Ambient Conditions on Optimal Supervisory Control. When the ratio of individual zone loads to total load does not change significantly with time, the optimal control variables are functions of the total sensible and latent gains to the zones and of the ambient dry- and wet-bulb temperatures. For systems with wet cooling towers and climates where moisture is removed from conditioned air, the effect of the ambient dry-bulb temperature alone is small because air enthalpy depends primarily on wet-bulb temperature, and the performance of wet-surface heat exchangers is driven primarily by the enthalpy difference. Typically, zone latent gains are on the order of 15 to 25% of the total zone gains, and changes in latent gains have a relatively small effect on performance for a given total load. Consequently, in many cases optimal supervisory control variables depend primarily on ambient wet-bulb temperature and total chilled-water load. However, load distributions between zones may also be important if they change significantly over time.
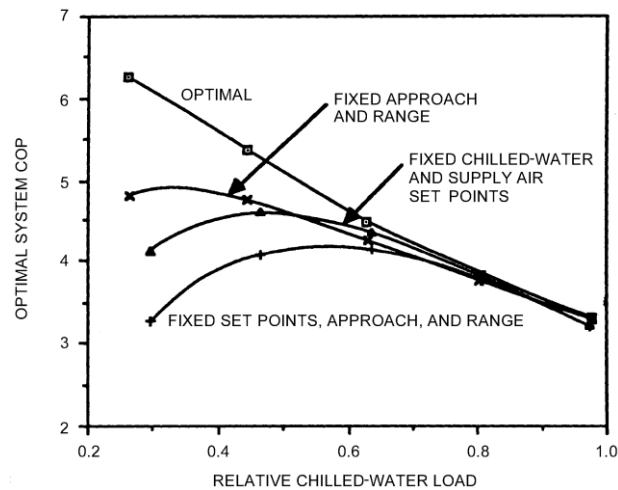
Generally, optimal chilled-water and supply air temperatures decrease with increasing load for a fixed ambient wet-bulb temperature and increase with increasing ambient wet-bulb temperature for a fixed load. Furthermore, optimal cooling tower airflow and condenser water flow rates increase with increasing load and ambient wet-bulb temperature.

Performance Comparisons for Supervisory Control Strategies

. Optimization of plant operation is most important when loads vary and when operation is far from design conditions for a significant period. Various strategies are used for chilled-water systems at off-design conditions. Commonly, the chilled-water and supply air set-point temperatures are changed only according to the ambient dry-bulb temperature. In some systems, cooling tower airflow and condenser water flow are not varied in response to changes in the load and ambient wet-bulb temperature. In other systems, these flow rates are controlled to maintain constant temperature differences between cooling tower outlet and ambient wet-bulb temperature (approach) and between cooling tower inlet and outlet (range), regardless of load and wet-bulb temperature. Although these strategies seem reasonable, they do not generally minimize operating costs.

Figure 12 shows a comparison of the COPs for optimal control and three alternative strategies as a function of load for a fixed ambient wet-bulb temperature. This system (from Braun et al. 1989b) incorporated the use of variable-speed pumps and fans. The three strategies are

Fig. 12   Comparisons of Optimal Control with Conventional Control Strategies

- Fixed chilled-water and supply air temperature set points (40 and 52°F [4.5 and 11°C], respectively), with optimal condenser-loop control
- Fixed tower approach and range (5 and 12°F [3 and 6.5 K], respectively), with optimal chilled-water loop control
- Fixed set points, approach, and range

Because the fixed values were chosen to be optimal at design conditions, the differences in performance for all strategies are minimal at high loads. However, at part-load conditions, Figure 12 shows that the savings associated with the use of optimal control can become significant. Optimal control of the chilled-water loop results in greater savings than that for the condenser loop for part-load ratios less than about 50%. The overall savings over a cooling season depend on the time variation of the load. If the cooling load is relatively constant and near the design load, fixed values of temperature set points, approach, and range could be chosen to give near-optimal performance. However, for typical building loads with significant daily and seasonal variations, the penalty for using a fixed set-point control strategy is typically in the range of 5 to 20% of the cooling system energy.

Even greater energy savings are possible with economizer control and discharge air temperature reset with constant-volume systems. Kao (1985) investigated the effect of different economizer and supply air reset strategies on both heating and cooling energy use for CAV, VAV, and dual-duct air handling systems for four different buildings. The results indicated that substantial improvements in a building's energy use may be obtained.

**Variable- Versus Fixed-Speed Equipment**
. Using variable-speed motors for chillers, fans, and pumps can significantly reduce energy costs but can also complicate the problem of determining optimal control. The overall savings from using variable-speed equipment over a cooling season depend on the time variation of the load. Typically, using variable-speed drives reduces equipment operating costs 20 to 50% compared to equipment with fixed-speed drives.

Figure 13 gives the overall optimal system performance for a cooling plant with either variable- or fixed-speed, variable-vane control of a centrifugal chiller. At part-load conditions, the system COP associated with using a variable-speed chiller is improved as much as 25%. However, the power requirements are similar at conditions associated with peak loads, because at full load the vanes are wide open and the speed under variable-speed control and fixed-speed operation is the same. The results in Figure 13 are from a single case study of a large chilled-water facility at the Dallas/Ft. Worth Airport (Braun et al. 1989b), constructed in the mid-1970s, where the existing chiller was retrofitted with a variable-speed drive. Differences in performance between variable- and fixed-speed chillers may be smaller for current equipment.
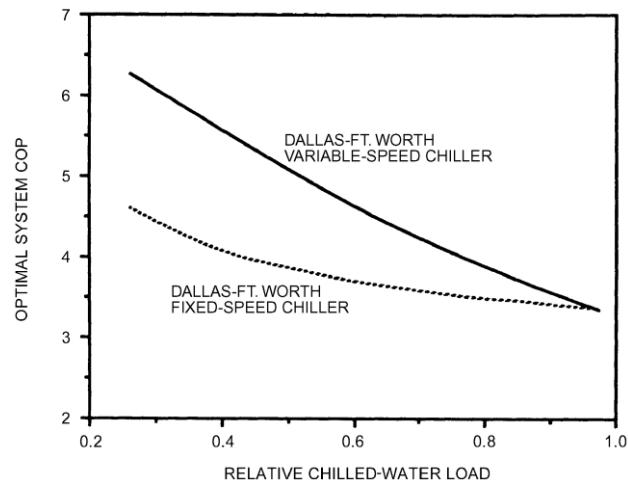
Fig. 13   Example of Optimal Performance for Variable- and Fixed-Speed Chillers

The most common design for cooling towers places multiple tower cells in parallel with a common sump. Each tower cell has a fan with one, two, or possibly three operating speeds. Although multiple cells with multiple fan settings offer wide flexibility in control, using variable-speed tower fans can provide additional improvements in overall system performance. Figure 14 shows an example comparison of optimal performance for single-speed, two-speed, and variable-speed tower fans as a function of load for a given wet-bulb temperature for a system with four cells (Braun et al. 1989b). The variable-speed option results in higher COP under all conditions. In contrast, for discrete fan control, the tower cells are isolated when their fans are off and the performance is poorer. Below about 70% of full-load conditions, there is a 15% difference in total energy consumption between single-speed and variable-speed fans. Between two-speed and variable-speed fans, the differences are much smaller, about 3 to 5% over the entire range.
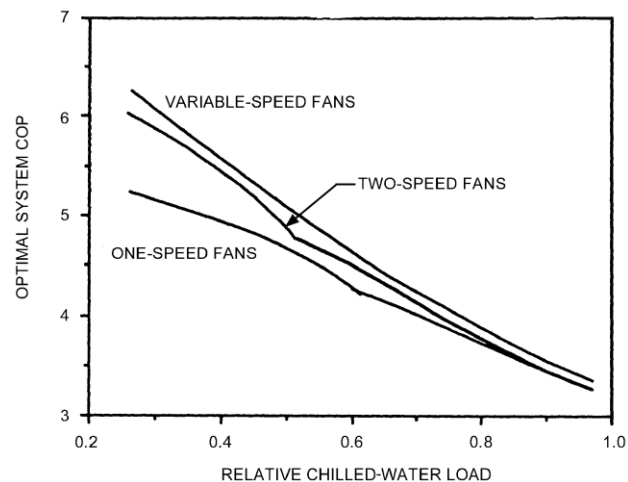
Fig. 14   Example Comparison of One-, Two-, and Variable-Speed Fans for Four-Cell Cooling Tower

Fixed-speed pumps that are sized to give proper flow to a chiller at design conditions are oversized for part-load conditions. Thus, the system will have higher operating costs than with a variable-speed pump of the same design capacity. Multiple pumps with different capacities have increased flexibility in control, and using a smaller fixed-speed pump for low loads can reduce overall power consumption. The optimal performance for variable-speed and fixed-speed pumps applied to both the condenser and chilled-water flow loops is shown in Figure 15 (Braun et al. 1989b). Large fixed-speed pumps were sized for design conditions; the small pumps were sized to have one-half the flow capacity of the large pumps. Below about 60% of full-load conditions, a variable-speed pump showed a significant improvement over the use of a single, large fixed-speed pump. With the addition of a small fixed-speed pump, improvements with the variable-speed pump were significant at about 40% of the maximum load.
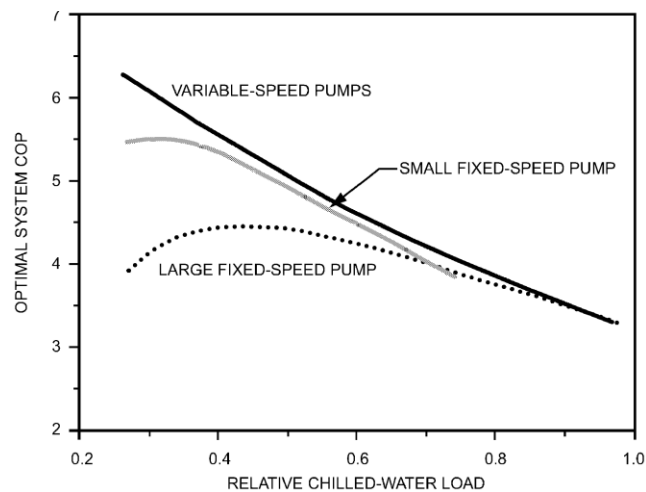
Fig. 15   Example of Optimal Performance for Variable- and Fixed-Speed Chillers

Fan energy consumed by VAV systems is strongly influenced by the device used to vary the airflow. Centrifugal fans with variable-speed drives typically provide the most energy-efficient performance. Brothers and Warren (1986) compared the fan energy consumption for a typical office building in various U.S. locations. The analysis focused on centrifugal and vaneaxial fans with three typical flow modulation devices: (1) dampers on the outlet side of the fan, (2) inlet vanes on the fan, and (3) variable-speed control of the fan motor. In all locations, the centrifugal fan used less energy than the vaneaxial fan. Vaneaxial fans have higher efficiencies at the full-load design point, but centrifugal fans have better off-design characteristics that lead to lower annual energy consumption. For a centrifugal fan, inlet vane control saved about 20% of the energy compared to damper control. Variable-speed control produced average savings of 57% compared to inlet vane control.

Hybrid Cooling Plants. Hybrid cooling plants use a combination of chillers that are powered by electricity and natural gas. Braun (2007a) developed a set of near-optimal operating strategies for hybrid cooling plants to reduce operating costs. Operating cost minimization for hybrid plants must account for effects of electrical and gas energy costs, electrical demand costs, and differences in maintenance costs associated with different chillers. Control strategies for hybrid cooling plants were developed by separating hourly energy cost minimization from the problem of determining trade-offs between monthly energy and demand costs. A demand constraint was set for each month, based on a heuristic strategy, and energy cost optimal strategies that attempted to satisfy the demand constraint were applied for cooling tower and chiller control at each decision interval. Simulated costs associated with the individual control strategies compared well with costs for optimal control.

Moreover, Braun (2007b) presented an algorithm for determining cooling tower fan settings in hybrid plants in response to loadings on individual chillers. Parameters of the algorithm were evaluated using design information for the chillers and cooling tower fans. In addition to reducing operating costs, use of the open-loop control strategy simplifies the control and improves the stability of tower control compared with using a constant condenser water supply or approach to wet-bulb. Simulated plant cooling costs associated with the algorithm were compared with costs for optimized settings, and were within 1% of the minimum costs. The developed control method is general, in the sense that it also applies to cooling plants that have all-electric or all-natural-gas chillers.

## DYNAMIC OPTIMIZATION FOR COOLING USING DISCRETE STORAGE

Thermal storage systems allow part, or all, of the cooling load to be shifted from on-peak to off-peak hours. Discrete cool storage systems typically use water or ice in unpressurized tanks as the thermal storage medium. Charging and discharging rates can be controlled from zero to a maximum design rate which in some cases may equal the design load. In building structure (passive) storage, the **building thermal mass** is the storage medium; charging and discharging are accomplished by adjusting space temperatures over a relatively narrow range. Thermally activated building systems (TABS), on the other hand, use structural elements with embedded pipes (e.g. rediant floor slabs) to promote controllable charging from within the mass. Conventional TABS relies on passive (uncontrolled) discharge. An active discharge mode has recently been demonstrated in a 250m2 (2700ft2) occupied building. Phase change materials may be used to increase the capacity of passive or TAB systems.

Utility incentives encouraging use of thermal storage are generally in the form of time-varying energy and peak demand charges. The commercial consumer is charged more for energy during the daytime and is also levied an additional charge each month based on the peak power consumption during the on-peak period. These incentives can be significant, depending on location, and are often the most important factor affecting an optimal control strategy for systems with thermal storage

Cooling Systems with Discrete Thermal Storage

Chapter 50 of the 2008 *ASHRAE Handbook—HVAC Systems and Equipment*, describes several possible storage media and system configurations for discrete thermal storage. Figure 16 depicts a generic storage system coupled to a cooling system and a building load.

The storage medium could be chilled water, ice or some other phase change material. Cooling equipment **charges** (operates at low temperatures to make ice) during unoccupied periods when the cost of electricity is low. During times of occupancy and higher electric rates, ice is melted (**storage discharging**) as the storage meets all or, in combination with the primary cooling equipment, part of the building load. Stratified water storage is charged and discharged from the bottom (cold end) of the tank with water returning from the load to the top of the tank during discharge and returning from the top of the tank to the chiller during charging.
.
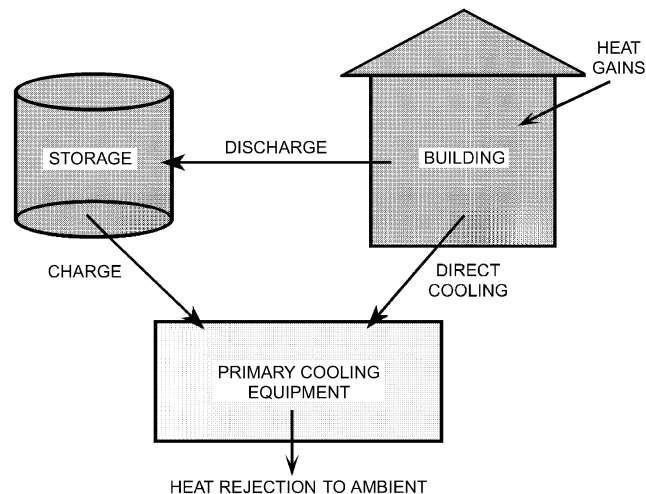
Fig. 16   Generic Storage System for Cooling.   Arrows Show direction of Heat Flow.

The primary control variables for the thermal storage systems depicted in Figure 16 are the rate of (1) energy removal from storage by the cooling system (charging rate) and (2) energy addition because of the load (discharging rate). Determining the optimal charging and discharging rates differs considerably from determining optimal set points for cooling plants that do not have storage. With thermal storage, control decisions (i.e., charging and discharging rates) determined for the current hour affect costs and control decisions for several hours in the future. Optimal control of thermal storage systems involves finding a sequence of charging and discharging rates that minimizes the total cost of providing cooling over an extended period of time, such as a day, and requires forecasting and application of dynamic optimization techniques. Constraints include limits on charging and discharging rates. The optimal control sequence results from tradeoffs between the costs of cooling the storage during off-peak hours and the cost of meeting the load during on-peak hours. In the absence of any utility incentives to use electricity at night, optimal control generally minimizes the use of ice storage because the cooling equipment operates less efficiently while charging at low temperatures. For building thermal mass systems, precooling increases heat gains from the ambient to the building. Close attention to system design and supervisory control is essential if this penalty is to be offset by potentially more efficient operation of the chiller at higher chilled-water and lower average heat rejection (night operation) temperatures.

Online optimal control of thermal storage, like other HVAC optimal control schemes, is rarely implemented because of the high initial costs associated with sensors (e.g., power) and software implementation. However, heuristic control strategies have been developed that provide near-optimal performance under most circumstances. The following sections provide background on developing control strategies for ice storage and building thermal mass. Detailed descriptions of some specific control strategies are given in the Supervisory Control Strategies and Tools section of this chapter.

Ice Storage. This section emphasizes ice storage applications, although much of the information applies to chilled-water storage as well. Figure 17 shows a schematic of a typical ice storage system. The system consists of one or more chillers, cooling tower cells, condenser water pumps, chilled-water/glycol distribution pumps, ice storage tanks, and valves for controlling charging and discharging modes of operation. Ice is made at night and used during the day to provide part of a building's cooling requirements; the storage is not sized to handle the full on-peak load requirement on the design day. Typically, in a load-leveling scheme, the storage and chiller capacity are sized such that chiller operates at full capacity during the on-peak period on the design day.

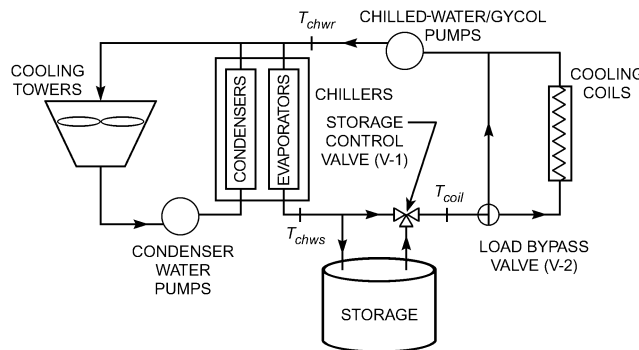Typical modes of operation for the system in Figure 17 are as follows:

Fig. 17 Schematic of an Ice Storage System

- **Storage charging mode:** Typically, charging (i.e., ice making) only occurs when the building is unoccupied and off-peak electric rates are in effect. In this mode, the load bypass valve V-2 is fully closed to the building cooling coils, the storage control valve V-1 is fully open to the ice storage tank (the total chilled-water/glycol flow is through the tank), and the chiller produces low temperatures (e.g., 20°F [–7°C]) sufficient to make ice within the tank.
- **Storage discharging mode:** Discharging of storage (i.e., ice melting) only occurs when the building is occupied. In this mode, valve V-2 is open to the building cooling coils and valve V-1 modulates the mixture of flows from the storage tank and chiller to maintain a constant supply temperature to the building cooling coils (e.g., 38°F [3°C]). Individual valves at the cooling coils modulate their chilled-water/glycol flow to maintain supply air temperatures to the zones.
- **Direct chiller mode:** The chiller may operate to meet the load directly without using storage during the occupied mode (typically when off-peak electric rates are in effect). In this mode, valve V-1 is fully closed with respect to the storage tank.

For a typical partial-storage system, the storage meets only a portion of the on-peak cooling loads on the design day and the chiller operates at capacity during the on-peak period. Thus, the peak power is limited by the capacity of the chiller. For off-design days, there are many different control strategies that meet the building's cooling requirements. However, each method has a different overall operating cost.

The best control strategy for a given day is a function of several factors, including utility rates, load profile, chiller characteristics, storage characteristics, and weather. For a utility rate structure that includes both time-of-use energy and demand charges, the optimal strategy can depend on variables that extend over a monthly time scale. Consider the charges typically associated with electrical use within a building. The first charge is the total cost of energy use for the building over the billing period, which is usually a month. Typically, the energy cost rate varies according to time of use, with high rates during the daytime on weekdays and low costs at night and on weekends. The second charge, the building demand cost, is the product of the peak power consumption during the billing period and the demand cost rate for that stage. The demand cost rate can also vary with time of day, with higher rates for on-peak periods. To determine a control strategy for charging and discharging storage that minimizes utility costs for a given system, it is necessary to perform a minimization of the total cost over the entire billing period because of the demand charge. An even more complicated cost optimization results if the utility rate includes ratchet clauses, whereby the demand charge is the maximum of the monthly peak demand cost and some fraction of the previous monthly peak demand cost within the cooling season. In either case, it is not worthwhile to perform an optimization over time periods longer than those for which reliable forecasts of cooling requirements or ambient conditions could be performed (e.g., 1 day). It is therefore important to have simple control strategies for charging and discharging storage over a daily cycle.

The following control strategies for limiting cases provide further insight:

- If the demand cost rate is zero and the energy cost rate does not vary with time, minimizing cost is equivalent to minimizing total electrical energy use. In general, cooling plant efficiency is lower when it is being used to make ice than when it is providing cooling for the building. Thus, in this case, the optimal strategy for minimum energy use minimizes the use of storage. Although this may seem like a trivial example, the most common control strategy in use today for partial ice storage systems, chiller-priority control, attempts to minimize the use of storage.
- If the demand cost rate is zero but energy costs are higher during on-peak than off-peak periods, minimizing cost then involves tradeoffs between energy use and energy cost rates. For relatively small differences between on-peak and off-peak rates of less than about 30%, energy penalties for ice making typically outweigh the effect of reduced rates, and chiller-priority control is optimal for many cases. However, with higher differentials between on-peak and off-peak energy rates or with chillers having smaller charging-mode energy penalties, the optimal strategy might maximize the use of storage. A control strategy that attempts to maximize the load-shifting potential of storage is called storage-priority control; in this scheme, the chiller operates during the off-peak period to fully charge storage. During the on-peak period, storage is used to cool the building in a manner that minimizes use of the chiller(s). Partial-storage systems that use storage-priority control strategies require forecasts for building cooling requirements to avoid prematurely depleting storage.
- If only on-peak demand costs are considered, then the optimal control strategy tends to maximize the use of storage and controls the discharge of storage in a manner to always minimize the peak building power. A storage-priority, demand-minimization control strategy for partial-storage systems requires both cooling load and noncooling electrical use forecasts.

A number of control strategies based on these three simple limiting cases have been proposed for ice storage systems (Braun 1992; Drees and Braun 1996; Grumman and Butkus 1988; Rawlings 1985; Spethmann 1989; Tamblyn 1985). Braun (1992) appears to have been the first to evaluate the performance of chiller-priority and storage-priority control strategies as compared with optimal control. The storage-priority strategy was termed load-limiting control because it attempts to minimize the peak cooling load during the on-peak period. For the system considered, the load-limiting strategy provided near-optimal control in terms of demand costs in all cases and worked well with respect to energy costs when time-of-day energy charges were available. However, the scope of the study was limited in terms of the systems considered.

Krarti et al. (1996) evaluated chiller-priority and storage-priority control strategies as compared with optimal control for a wide range of systems, utility rate structures, and operating conditions. Similar to Braun (1992), they concluded that load-limiting, storage-priority control provides near-optimal performance when there are significant differentials between on-peak and off-peak energy and demand charges. However, optimal control provides superior performance in the absence of time-of-day incentives. In general, the monthly utility costs associated with chiller-priority control were significantly higher than optimal and storage-priority control. However, without time-of-use energy charges, chiller-priority control did provide good performance for individual days when the daily peak power was less than the monthly peak. General guidance based on the work is presented by Henze et al. (2003). Drees and Braun (1996) developed a simple rule-based control strategy that combines elements of storage- and chiller-priority strategies in a way that results in near-optimal performance under all conditions. The strategy was derived from heuristics obtained through both daily and monthly optimization results for several simulated systems. A modified version of this strategy is presented in the Supervisory Control Strategies and Tools section of this chapter.

Braun (2007c, 2007d) also developed a near-optimal control method for charging and discharging of cool storage systems when RTP electric rates are available The algorithm requires relatively low-cost measurements (cooling load and storage state of charge) and very little plant specific information, is computationally simple, and ensures that building cooling requirements are always met (e.g., storage is not prematurely depleted). The control method was evaluated for ice storage systems using a simulation tool for different combinations of cooling plants, storage sizes, buildings, locations, and RTP rates.

Control Strategies for Cooling Systems with Discrete Thermal Storage

The choice of a control strategy for a cooling system with discrete thermal storage system results from a tradeoff between performance (i.e., operating costs) and ease of implementation (i.e., initial costs). Chiller-priority control has the lowest implementation costs, but generally leads to the highest operating costs. Storage-priority strategies provide superior performance, but require the use of a forecaster and a measurement of state of charge for storage. This section presents details of chiller-priority, load-limiting, and rule-based control applied to ice storage systems. Each of these strategies shares the same procedures for charging storage, but differs in the manner in which storage is discharged. In general, the control strategies presented in this section are appropriate for systems with utility rate structures that include time-of-use energy and demand charges, but would not be appropriate in conjunction with real-time pricing. Additional information on control strategies for cool storage systems can be found in Chapter 50 of the 2008 *ASHRAE Handbook—HVAC Systems and Equipment*.

Charging Strategies

Ice making should be initiated when both the building is unoccupied and off-peak electrical rates are in effect. During the ice-making period, the chiller should operate at full capacity. Cooling plants for ice storage generally operate most efficiently at full load because of the auxiliaries and the characteristics of ice-making chillers. With feedback control of the chilled-water/glycol supply temperature, full capacity control is accomplished by establishing a low enough set point to ensure this condition (e.g., 20°F [–7°C]).

Internal Melt Storage Tanks. The chiller should operate until the tank reaches its maximum state of charge or the charging period (i.e., off-peak, unoccupied period) ends. This strategy ensures that sufficient ice will be available for the next day without the need for a forecaster. Typically, only a small heat transfer penalty is associated with restoring a partially discharged, internal melt storage tank to a full charge. For this type of storage device, the charging cycle always starts with a high transfer effectiveness because water surrounds the tubes regardless of the amount of ice melted. The heat transfer effectiveness drops gradually until the new ice formations intersect with old formations, at which point the tank is fully re-charged.

External Melt Storage Tanks. These tanks have a more significant heat transfer penalty associated with recharging after a partial discharge, because ice forms on the outside of existing formations during charging. In this case, it is more efficient to fully discharge the tank each day and only recharge as necessary to meet the next day's cooling requirements. To ensure that adequate ice is available, the maximum possible storage capacity needed for the next day must be forecast. The storage requirements for the next day depend on the discharge strategy used and the building load. In general, the state of charge for storage necessary to meet the next day's load can be estimated according to

$$X_{chg} = \sum_{k=1}^{\substack{\text{occupied} \\ \text{period}}} \frac{\hat{Q}_{load,k} - \hat{Q}_{ch,k}}{C_s} \tag{54}$$

where X_{chg} is the relative state of charge at the end of the charging period, C_s is the maximum change in internal energy of the storage tank that can occur during a normal discharge cycle, and $\hat{Q}_{load,k}$ and $\hat{Q}_{ch,k}$ are forecasts of the building load and chiller cooling requirement for the *k*th stage (e.g., hour) of the occupied period. The relative state of charge is defined in terms of two reference states: the fully discharged and fully charged conditions that correspond to values of zero and one. These conditions are defined for a given storage based on its particular operating strategy (ASHRAE *Standard* 150; Elleson 1996). The fully charged condition exists when the control stops the charge cycle as part of its normal sequence. Similarly, the fully discharged condition is the point where no more usable cooling is recovered from the tank. Typically, zero state of charge corresponds to a tank of water at a uniform temperature of 32°F (0°C)

DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*

and a complete charge is associated with a tank having maximum ice build at 32°F (0°C). (The fully discharged and fully charged conditions are arbitrarily selected reference states.) In abnormal circumstances, a storage tank can be discharged or charged beyond these conditions, resulting in relative states of charge below zero or above one.

Hourly forecasts of the next day's cooling requirement can be determined using the algorithm described in the section on Forecasting Diurnal Energy Requirements. However, long-term forecasts are highly uncertain and a safety factor based on previous forecast errors is appropriate (e.g., uncertainty of two or three times the standard deviation of the errors of previous forecasts). Estimates of hourly chiller requirements should be determined using the intended discharge strategy (described in the next section) and building load forecasts.

Discharging Strategies

Three discharge strategies are presented for use with utility structures having on-peak and off-peak energy and demand charges: (1) chiller-priority control, (2) storage-priority, load-limiting control, and (3) a rule-based strategy that uses both chiller-priority and load-limiting strategies.

Chiller-Priority Discharge. During the storage discharge mode, the chiller operates at full cooling capacity (or less if sufficient to meet the load) and storage matches the difference between the building requirement and chiller capacity. For the example system illustrated in Figure 17, the chiller supply temperature set point t_{chws} is set equal to the desired supply temperature for the coils t_{coil}. If the capacity of the chiller is sufficient to maintain this set point, then storage is not used and the system operates in the direct chiller mode. Otherwise, the storage control valve modulates the flow through storage to maintain the supply set point, providing a cooling rate that matches the difference between the building load and the maximum cooling capacity of the chillers.

This strategy is easy to implement and does not require a load forecast. It works well for design conditions, but can result in relatively high demand and energy costs for off-design conditions because the chiller operates at full capacity during the on-peak period.

Storage-Priority, Load-Limiting Control. Several storage-priority control approaches ensure that storage is not depleted prematurely. Braun (1992) presented a storage-priority control strategy, termed *load-limiting* control, which tends to minimize the peak cooling plant power demand. The operation of equipment for load-limiting control during different parts of the occupied period can be described as follows:

- **Off-Peak, Occupied Period.** During this period, the goal is to minimize use of storage, and the chiller-priority described in the previous section should be applied.
- **On-Peak, Occupied Period.** During this period, the goal is to operate the chillers at a constant load while discharging the ice storage such that the ice is completely melted when the off-peak period begins. This requires using a building cooling-load forecaster. At each decision interval (e.g., 15 min), the following steps are applied:

  (1) Forecast the total integrated building cooling requirement until the end of the discharging period.
  (2) Estimate the state of charge of the ice storage tank from measurements.
  (3) At any time, the chiller loading for load-limiting control is determined as

  $$\dot{Q}_{LLC} = \text{Max}\left[\frac{\dot{Q}_{load,occ} - (X - X_{min})C_s}{\Delta t_{on}}, \dot{Q}_{ch,min}\right] \tag{55}$$

  where $\dot{Q}_{load,occ}$ is a forecast of the integrated building load for the rest of the on-peak period, Δt_{on} is the time remaining in the on-peak period, X is the current state of charge defined as the fraction of the maximum storage capacity, X_{min} is a minimum allowable state of charge, C_s is the maximum possible energy that could be added to storage during discharge, and $\dot{Q}_{ch,min}$ is the minimum allowable chiller cooling capacity. If the chiller does not need to be operated during the remainder of the occupied, on-peak period, the minimum allowable cooling capacity could be set to zero. Otherwise, the cooling capacity should be set to the minimum at which the chiller can safely operate.

  (4) Determine the chiller set-point temperature necessary to achieve the desired loading as

  $$t_{chws} = t_{chwr} - \frac{\dot{Q}_{LLC}}{C_{chw}} \tag{56}$$

  where t_{chwr} is the temperature of water/glycol returned to the chiller and C_{chw} is the capacitance rate (mass flow times specific heat) of the flow stream.

Hourly forecasts of cooling loads can be determined using the algorithm described in the section on Forecasting Diurnal Energy Requirements. The hourly forecasts are then integrated to give a forecast of the total cooling requirement. To ensure sufficient cooling capacity, a worst-case forecast of cooling requirements could be estimated as the sum of the best forecast and two or three times the standard deviation of the errors of previous forecasts.

Rule-Based Controller. Drees and Braun (1996) presented a rule-based controller that combines elements of chiller-priority and storage-priority strategies, along with a demand-limiting algorithm to achieve near-optimal control. The demand-limiting algorithm requires a measurement of the total building electrical use. A simpler strategy is described here that does not require this measurement and yields equivalent performance whenever the peak demand for the billing period is coincident with the peak cooling load.
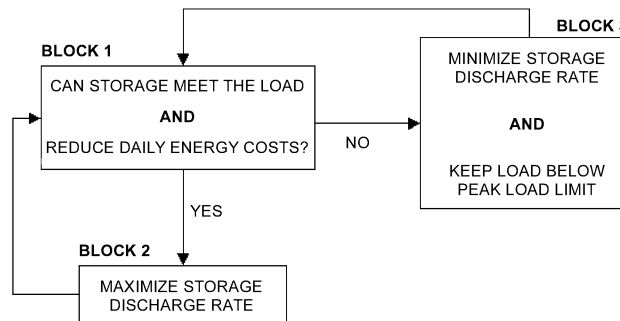
Fig. 41   Flowchart for Rule-Based Controller Discharge Strategy

Figure 41 shows a flowchart for the discharge strategy that is applied during each decision interval (e.g., 15 min) during the occupied period. Block 1 determines whether storage use should be maximized or minimized. Block 2 is used if storage use lowers daily energy costs and storage is sufficient to meet the remainder of the load for the occupied period without operating the chillers. Otherwise, the goal of the strategy in block 3 is to minimize storage use while keeping peak load below a limit. This strategy tends to keep the chiller(s) heavily loaded (and therefore, if part-load efficiency is poor, operating efficiently) until they are no longer needed. The logic in each block is as follows:

- **Block 1: Discharge Strategy Selection.** The discharge of storage will not reduce the energy cost whenever the cost of replenishing the ice is greater than the cost of providing direct cooling by the chiller(s). This situation is always the case during the off-peak, occupied period because the electricity rates are the same as those associated with the charging period and chillers are less efficient in ice-making mode than when providing direct cooling. Furthermore, during the on-peak, occupied period, using storage generally reduces energy costs whenever the following criterion holds:

$$ECR > COP_d/COP_c \qquad (57)$$

where ECR is the ratio of on-peak to off-peak energy charges and COP_d and COP_c are coefficients of performance for the cooling plant (including chiller, pumps, and cooling tower fans) during discharging and charging of the tank. The COPs should be evaluated at the worst-case charging and discharging conditions associated with the design day. Typically, this ratio is between about 1.2 and 1.8 for systems with cooling towers. However, the ratio can be lower because of the effect of cool nighttime temperatures, especially for systems with air-cooled condensers in dry climates.

If the criterion of Equation (57) is satisfied, control switches from chiller-priority to storage-priority strategy whenever storage capacity is greater than the remaining integrated load. Therefore, storage-priority control is enabled whenever

$$(X - X_{min})C_s \geq \dot{Q}_{load, occ} \qquad (58)$$

Hourly forecasts of cooling loads can be determined using the algorithm described in the section on Forecasting Diurnal Energy Requirements and then integrated to give a forecast of the total cooling requirement. To ensure that adequate ice is available, worst-case hourly forecasts can be determined by adding the expected value of the hourly forecasts and the forecast errors associated with a specified confidence interval (e.g., two standard deviations for a 95% confidence interval). The worst-case hourly forecasts can then be integrated to give a worst-case integrated forecast.

- **Block 2: Maximum Use of Storage.** In this mode, the chillers are turned off and storage is used to meet the entire load throughout the remainder of the occupied period. However, a chiller may need to be turned on if the storage discharge rate is not sufficient to meet the building load (i.e., the coil supply temperature set point cannot be maintained).

- **Block 3: Minimize Use of Storage with Peak Load Limiting.** At any time, a target chiller load is determined as

$$\dot{Q}_{ch} = Min[Max(\dot{Q}_{ch,peak}, \dot{Q}_{LCC}), \dot{Q}_{load}] \qquad (59)$$

where $\dot{Q}_{ch,peak}$ is the peak chiller cooling requirement that has occurred during the on-peak period for the current billing period, $\dot{Q}_{LCC}$ is the chiller load associated with load-limiting control and determined with Equation (55), and $\dot{Q}_{load}$ is the current building load. The chiller set-point temperature necessary to achieve the desired loading is determined as

$$t_{chws} = Max\left(t_{chwr} - \frac{\dot{Q}_{ch}}{C_{chw}}, t_{coil}\right) \qquad (60)$$

On the first day of each billing period, $\dot{Q}_{ch,peak}$ is set to zero. For this first day, applying Equation (59) leads to the load-limiting control strategy described in the previous section. On subsequent days, load-limiting control is used only if the current peak limit would lead to premature depletion of storage. Whenever the current load is less than $\dot{Q}_{ch,peak}$ and $\dot{Q}_{LCC}$, Equations (59) and (60) lead to chiller-priority control.

DYNAMIC OPTIMIZATION FOR COOLING USING THERMAL MASS OR TABS

Dynamic Optimization for Precooling of Building Thermal Mass

For conventional night setup strategies, it is well known (Bloomfield and Fisk 1977) that building mass works to increase operating costs. A massless building would require no time for precooling or preheating and would have lower overall cooling or heating loads than an actual building. However, under proper circumstances, using a building's thermal storage for load shifting can significantly reduce operational costs and energy use, even though the total zone loads may increase. This is especially true for cooling of high-performance buildings, in which internal loads may dominate.

At any given time, the cooling requirement for a space is caused by convection from internal gains (lights, equipment, and people) and interior surfaces. Because a significant fraction of the internal gain is radiated to interior surfaces, the state of a building's thermal storage and the convective coupling dictates the cooling requirement. Precooling the building during unoccupied times reduces the overall convection from exposed surfaces during the occupied period as compared with night setup control and can reduce daytime cooling requirements. The potential for storing thermal energy in the structure and furnishings of conventional commercial buildings is significant when compared to the load requirements. Typically, internal gains are about 2 to 4 W per square foot (20 to 40 W per square meter) of floor space. The thermal capacity for typical concrete building structures is approximately 2 to 4 W·h/°F per square foot [40 to 80 (W·h)/K per square meter] of floor area. Thus, for an internal space, the energy storage can handle the load for about 1 h for every 1°F (0.6 K) of precooling of the thermal mass.

Opportunities for reducing operating costs by using building thermal mass for cooling derive from four effects: (1) reduction in demand costs, (2) use of low-cost off-peak electrical energy, (3) reduced mechanical cooling from the use of cool nighttime air for ventilation precooling, and (4) improved mechanical cooling efficiency from increased operation at more favorable part-load and ambient conditions. However, these benefits must be balanced with the increase in total cooling requirement that occurs with precooling the thermal mass. Therefore, the savings associated with load shifting and demand reductions depend on both the method of control and the specific application.

Several simulation studies have been performed that demonstrate a substantial benefit to precooling buildings in terms of cost savings and peak cooling load reduction (Andresen and Brandemuehl 1992; Braun 1990; Rabl and Norford 1991; Snyder and Newell 1990). Possible energy savings ranged from 0 to 25%; possible reductions in total building peak electrical demand ranged from 15 to 50% compared with conventional control. The results can be sensitive to the convective coupling between the air and the thermal mass, and the mass of the furnishings may be important (Andresen and Brandemuehl 1992).

Determining the optimal set of building temperatures over time that minimizes operating costs is complex. Keeney and Braun (1996) developed a simplified approach for determining optimal control of building thermal mass using two optimization variables for the precool period and a set of rules for the occupied period of each day. This approach significantly reduces the computation required for determining the optimal control as compared with considering hourly zone set points as optimization variables. Results of the simplified approach compared well with those of detailed optimizations for a range of systems (over 1000 different combinations of building types, weather conditions, cooling plants, and utility rates).

Morris et al. (1994) performed a set of experiments using a test facility at the National Institute of Standards and Technology (NIST) to demonstrate the potential for load shifting and load leveling when control was optimized. Two different control strategies were considered: (1) minimum cooling system energy use and (2) minimum peak cooling system electrical demand. The two strategies were implemented in the test facility and compared with night setup control. Figure 18 shows the 24 h time variation in the cooling requirement for the test facility allowed to reach a steady-periodic condition for both the minimum energy use strategy and conventional night setup control. The results indicate a significant load-shifting potential for the optimal control. Overall, the cooling requirements during the occupied period were approximately 40% less for optimal than for night setup control.

Comfort conditions were also monitored for the tests. Figure 19 gives the time variation of predicted mean vote (PMV) for the two control strategies as determined from measurements at the facility. A PMV of zero is a thermally neutral sensation, positive is too warm, and negative too cool. In the region of ±0.5, comfort is not compromised to any significant extent. Figure 19 shows that the comfort conditions were essentially identical for the two control methods during the occupied period. The space temperature, which has the dominant effect on comfort, was maintained at 75°F (24°C) during the occupied period for both control methods. During the unoccupied period, the cooling system was off for night setup control and the temperature floated to warm comfort conditions. On the other hand, the optimal controller precooled the space, resulting in cool comfort conditions before occupancy. During these tests, the minimum space temperature during precooling was 68°F (20°C), and the space temperature set point was raised to 75°F (24°C) just before occupancy.
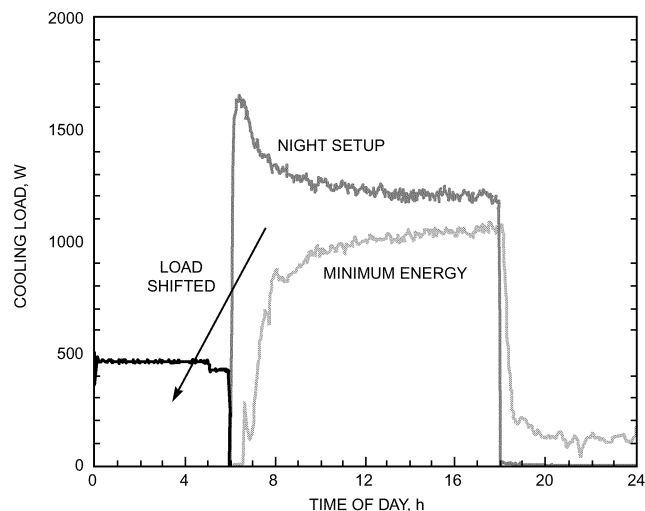
Fig. 18   Comparison of Cooling Requirements for Minimum Energy and Night Setup Control
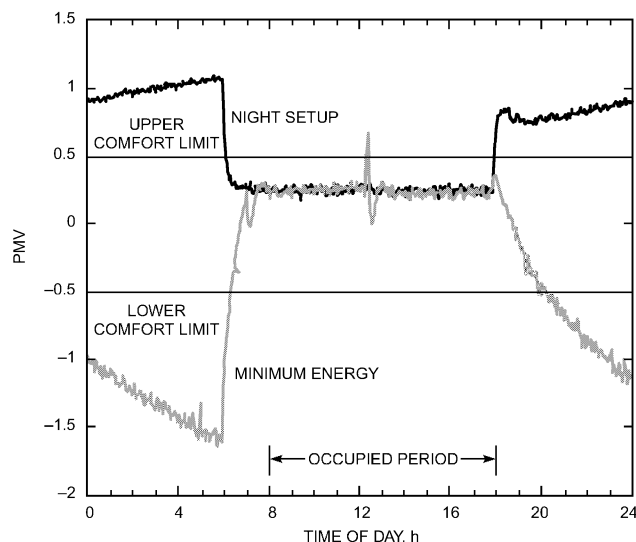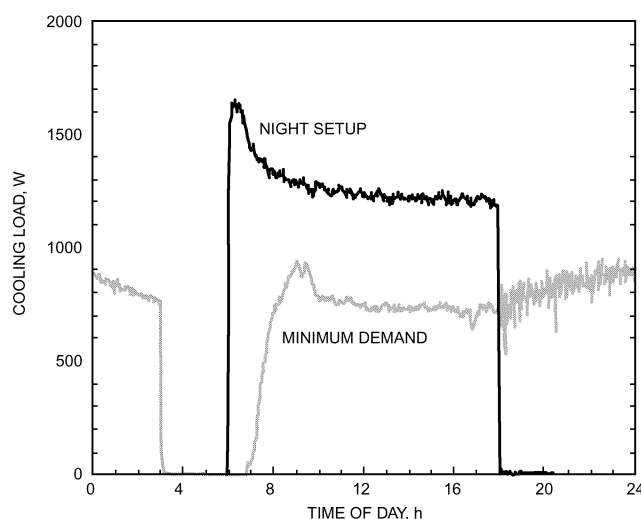
(Morris et al. 1994)

Fig. 19   Comparison of Predicted Mean Vote (PMV) for Minimum Energy and Night Setup Control

(Morris et al. 1994)

Figure 20 shows the 24 h time variation in the cooling requirement for the test facility for both the minimum peak demand strategy and conventional night setup. Optimal control involved precooling the structure and adjusting the space temperatures within the comfort zone ($-0.5 < \text{PMV} < 0.5$) during the occupied period to achieve the minimum demand. Although the true minimum was not achieved during the tests, the peak cooling rate during the occupied period was approximately 40% less for minimum peak demand control than for night setup control.

Morris et al. (1994) demonstrated significant savings potential for control of building thermal mass; however, they also showed that the cost savings are very sensitive to the application, operating conditions, and method of control. For example, an investigation into the effect of precooling on the on-peak cooling requirements for an existing building (which may not have been a good candidate for use of building thermal storage) showed only a 10% reduction in the cooling energy required during the occupied period, with a substantial increase in the total cooling required and no reduction in the peak cooling requirement (Ruud et al. 1990). System simulations can be used to identify (1) whether the system is a good candidate for using building thermal mass and (2) an effective method for control, before implementing a strategy in a particular building.

Fig. 20   Comparison of Cooling Requirements for Minimum Demand and Night Setup Control

(Morris et al. 1994)

Table 2   Cooling Season Energy, Demand, and Total Costs and Savings Potential of Different Building Mass Control Strategies

Strategy	Costs in U.S. Dollars			Savings, %
	Energy	Demand	Total	
Night setup	$ 90,802	$ 189,034	$ 279,836	0.0
Light precool	$ 84,346	$ 147,581	$ 231,928	17.1
Moderate precool	$ 83,541	$ 143,859	$ 227,400	18.7
Extended precool	$ 81,715	$ 134,551	$ 216,266	22.7
Maximum discharge	$ 72,638	$ 91,282	$ 163,920	41.4
Two-hour linear rise	$ 72,671	$ 91,372	$ 164,043	41.4
Four-hour linear rise	$ 73,779	$ 115,137	$ 188,916	32.5
Nine-hour linear rise	$ 77,095	$ 141,124	$ 218,219	22.0

Source: Braun et al. (2001).
Note: Building located in Chicago, Illinois.

Table 2   Cooling Season Energy, Demand, and Total Costs and Savings Potential of Different Building Mass Control Strategies

Strategy	Costs in U.S. Dollars			Savings, %
	Energy	Demand	Total	
Night setup	$90 802	$189 034	$279 836	0.0
Light precool	$84 346	$147 581	$231 928	17.1
Moderate precool	$83 541	$143 859	$227 400	18.7
Extended precool	$81 715	$134 551	$216 266	22.7
Maximum discharge	$72 638	$91 282	$163 920	41.4
Two-hour linear rise	$72 671	$91 372	$164 043	41.4
Four-hour linear rise	$73 779	$115 137	$188 916	32.5
Nine-hour linear rise	$77 095	$141 124	$218 219	22.0

Source: Braun et al. (2001).
Note: Building located in Chicago, Illinois.

Keeney and Braun (1997) used system simulation to develop a control strategy that was then tested in a large commercial building located northwest of Chicago. The goal of the control strategy was to use building thermal mass to limit the peak cooling load for continued building operation in the event of the loss of one of the four central chiller units. The algorithm was tested using two nearly identical buildings separated by a large, separately cooled entrance area. The east building used the existing building control strategy; the west building used the precooling strategy developed for this project. The precooling control strategy successfully limited the peak load to 75% of the cooling capacity for the west building, whereas the east building operated at 100% of capacity. Details of the strategy and case study results are presented in the Supervisory Control Strategies and Tools section of this chapter.

Braun et al. (2001) used on-site measurements from the same building used by Keeney and Braun (1997) to train site-specific models that were then used to develop site-specific control strategies for using building thermal mass and to evaluate the possible cost savings of

Applications I-P and (SI)   ©2015 ASHRAE, Inc.

DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*

these strategies. The building is an excellent candidate for using building thermal mass because it has (1) a large differential between on-peak and off-peak energy rates (about a 2-to-1 ratio), (2) a large demand charge (about $16/kW), (3) a heavy structure with significant exposed mass, and (4) cooling loads that are dominated by internal gains, leading to a high storage efficiency. The model underpredicted the total HVAC bill by about 5% but worked well enough to be used in comparing the performance of alternative control strategies.

Table 2 gives estimates of cooling-related costs and savings over the course of three summer months for different control strategies. The light precool and moderate precool strategies are simple strategies that precool the building at a fixed set point of 67°F (19.5°C) before occupancy and then maintain a fixed discharge set point in the middle of the comfort range (73°F [23°C]) during occupancy. The light precool begins at 3 AM, whereas moderate precool starts at 1 AM. The extended precool strategy attempts to maintain the cooled thermal mass until the onset of the on-peak period. In this case, the set point at occupancy is maintained at the lower limit of comfort (69°F [20.5°C]) until the on-peak period begins at 10 AM. At this point, the set point is raised to the middle of the comfort range (73°F [23°C]). The other strategies use the extended precooling, but the entire comfort range is used throughout the on-peak, occupied period. The maximum discharge strategy attempts to discharge the mass as quickly as possible after the on-peak period begins. In this case, the set point is raised to the upper limit of comfort within an hour after the on-peak period begins. The slow linear rise strategy raises the set point linearly over the entire on-peak, occupied period (9 h in this case), whereas the fast linear rise strategy raises the set point over 4 h.

The strategies that do not use the entire comfort range during the occupied period (light precool, moderate precool and extended precool) all provided about 20% savings compared to night setup. Each of these strategies reduced both energy and demand costs, but the demand costs and reductions were significantly greater than the energy costs and savings. The decreases in energy costs were due to favorable on-to-off peak energy rate ratios of about 2 to 1. The high on-peak demand charges provided even greater incentives for precooling. The savings increased with the length of the precooling period, particularly when precooling was performed close to the onset of on-peak rates. The maximum discharge strategy, which maximizes discharge of the thermal storage within the structure, provided the largest savings (41%). Much of the additional savings came from reduced demand costs. The linear rise strategies also provided considerable savings with greater savings associated with faster increases in the set point temperature.

Morgan and Krarti (2006) performed both simulation analyses and field testing to evaluate various precooling strategies. They found that energy cost savings associated with precooling thermal mass depends on several factors, including thermal mass level, climate, and utility rate. For time-of-use (TOU) utility rates, they found that energy cost savings are primarily affected by the ratio of on-peak to off-peak demand charges as well as the ratio of on-peak to off-peak energy charges (refer to Chapter 50 of the 2008 *ASHRAE Handbook—HVAC Systems and Equipment*).

More recently, an extensive study of building thermal mass control was conducted in the context of ASHRAE research project RP-1313 (Henze et al. 2007), in which optimal building thermal mass control strategies were investigated for time-of-use electric utility rates structures, including demand charges, with the help of a newly developed integrated optimization and building simulation tool (Henze et al. 2008). Cheng et al. (2008) identified the primary factors that influence optimal control of passive thermal storage, where optimal control strategies are determined with the objective of minimizing total energy and demand costs. A fractional factorial analysis was used to investigate how cost savings are affected by several building and system characteristics, utility rate structures, and climates. Utility rates, internal loads, building mass level, and equipment efficiency were found to have the largest impacts on cost savings, whereas building envelope characteristics did not have a significant impact. Although the magnitude of savings is affected by climate, the relative impacts of each of these factors are largely independent of weather.

Using the same simulation and optimization environment, Henze et al. (2009) presented advances toward near-optimal building thermal mass control derived from full factorial analyses of the important parameters influencing passive thermal storage for a range of buildings and climate/utility rate structure combinations. In response to the actual utility rates imposed in the investigated cities, insights and control simplifications were derived from those buildings deemed suitable candidates. The near-optimal strategies were derived from the optimal control trajectory, consisting of four variables, and then tested for effectiveness and validated with respect to uncertainty regarding building parameters and climate variations. Although no universally applicable control guideline could be found, a significant number of cases (i.e., combinations of buildings, weather, and utility rate structure) were investigated and offer both insight into and recommendations for simplified control strategies. These recommendations are a good starting point for experimentation with building thermal mass control for a substantial range of building types, equipment, climates, and utility rates.

The cost savings potential of optimal passive thermal storage controls were examined by Greensfelder et al. (2011) for the case of day-ahead, real-time electricity rate structures. The operational strategies of three office building models were optimized in four U.S. cities (Chicago, New York, Houston, and Los Angeles) using price and weather data for the summer of 2008. Building thermal mass was optimized using a predictive optimal controller to define supervisory control strategies in terms of building global cooling temperature set points. A global minimization algorithm determined optimal set-point trajectories for each day divided into four distinct time periods (called building modes). Cost savings were found to range from 0 to 14%, depending on the building, climate, and characteristics of the rate signal. The best cost savings occurred in the presence of price spikes or cool nighttime temperatures. Moreover, it was found that low internal gains favored a more flexible precooling strategy, whereas high internal gains coupled with low thermal mass resulted in poor precooling performance.

Dynamic Optimization for Thermally Activated Building Systems (TABS)

One way to reduce cooling energy further along the path to net-zero-energy buildings is **thermally activated building systems (TABS):** direct cooling of building mass by chilled water in conjunction with chillers designed for very high part-load efficiency in **low-lift** operation and enthalpy-recovery dedicated outdoor air systems (DOAS).

TABS differs from passive storage by cooling the mass directly from the inside, thus eliminating charging-mode convective or radiative coupling resistances. By directly precooling the mass, instead of the occupied space, substantially better storage efficiency and larger effective diurnal storage capacity are achieved. Precooling energy percentage savings may be 5 to 35% higher because the condenser-evaporator temperature difference is lower to begin with and because, with the elimination of supply fans and mechanical cooling, energy

use is dominated by compressor operation at very low pressure ratios (Armstrong et al. 2009; Katipamula et al. 2010a). Economizer-mode, which involves mainly pump rather than fan transport energy, is extremely energy efficient as well.

Katipamula et al. (2010) simulated idealized thermal energy storage (TES) to approximate TABS, and a variable-speed air-cooled chiller with variable-speed distribution pump. The TES and distribution options are shown schematically in Figure 21.
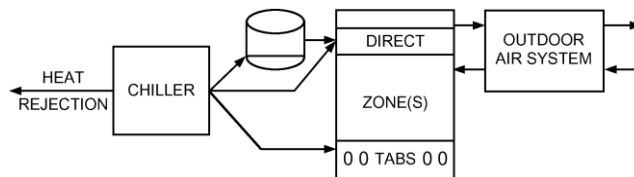
Fig. 21   Schematic of a Thermally Activated Building System with Three Direct Cooling Options

The chiller/distribution system was modeled with compressor, condenser fan, and chilled-water pump capable of wide 10:1 speed ranges and statically optimized control, resulting in evaporator temperatures of up to 64°F (18°C) at 10% of design load. Chiller performance was measured by an independent testing laboratory over a wide range of part-load fraction, outdoor temperatures, and chilled-water temperatures (Katipamula et al. 2010b). Performance curves fit to the data are shown in Figure 22. The performance map on the left is for a standard chilled-water reset schedule recommended for VAV systems (ASHRAE *Standard* 90.1), and the right-hand map represents performance of the chiller with radiant ceiling panels (RCPs) or chilled-beam distribution system serving a conditioned space at T_{op} = 75°F (24°C). Compressor, condenser fan, and chilled-water pump are operated at optimal speed under any given condition and part-load fraction.

A flat electricity rate was assumed, to achieve minimum annual energy use, and the performance of standard and high-performance versions of 12 building types was simulated in16 climates. The load shifting achieved by TES is illustrated in Figure 23. The joint distribution of chiller output is expressed in full-load operating hours for each bin cell of 5°F (2.8 K) in outdoor temperature by 10% in rated capacity. Compared to the base-case load distribution, operation hours for the TES system are significantly shifted in two respects. The bin in which the full-load equivalent operating hours (FLEOH) peak occurs is typically 15°F (8.3 K) lower than in the baseline chiller load distribution. Although the chiller continues to operate at high outdoor temperatures, it does so at much lower part-load ratios. Conversely, the FLEOH of operation at low outdoor temperature and low part-load ratio significantly increased.
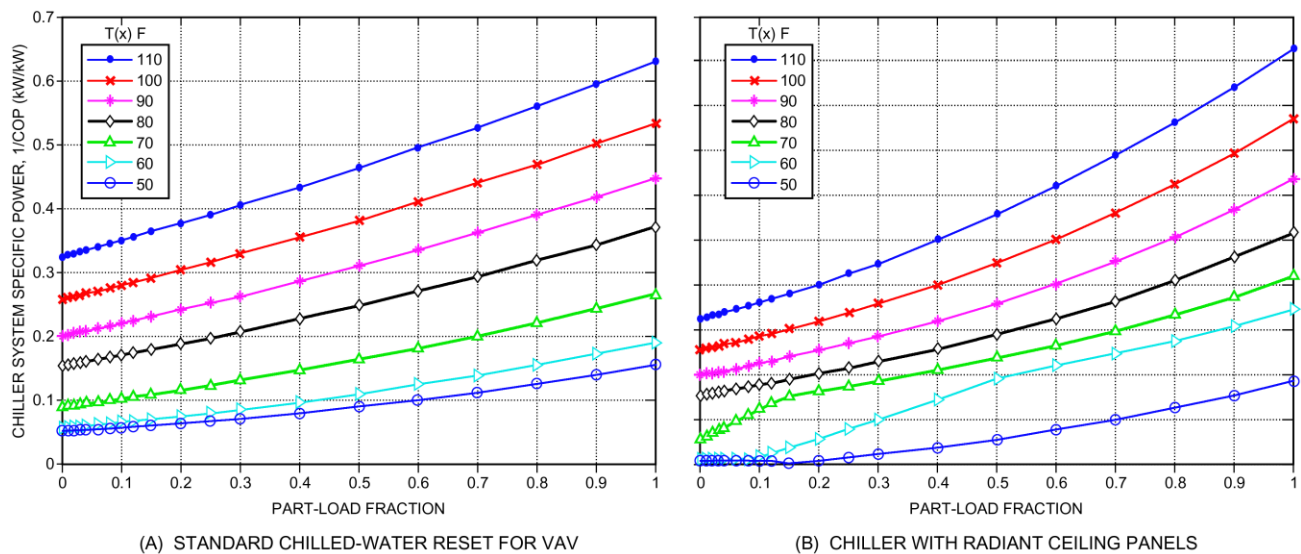
Fig. 22   Performance of an Optimally Controlled Chiller for Two Different Load-Side Boundary Conditions
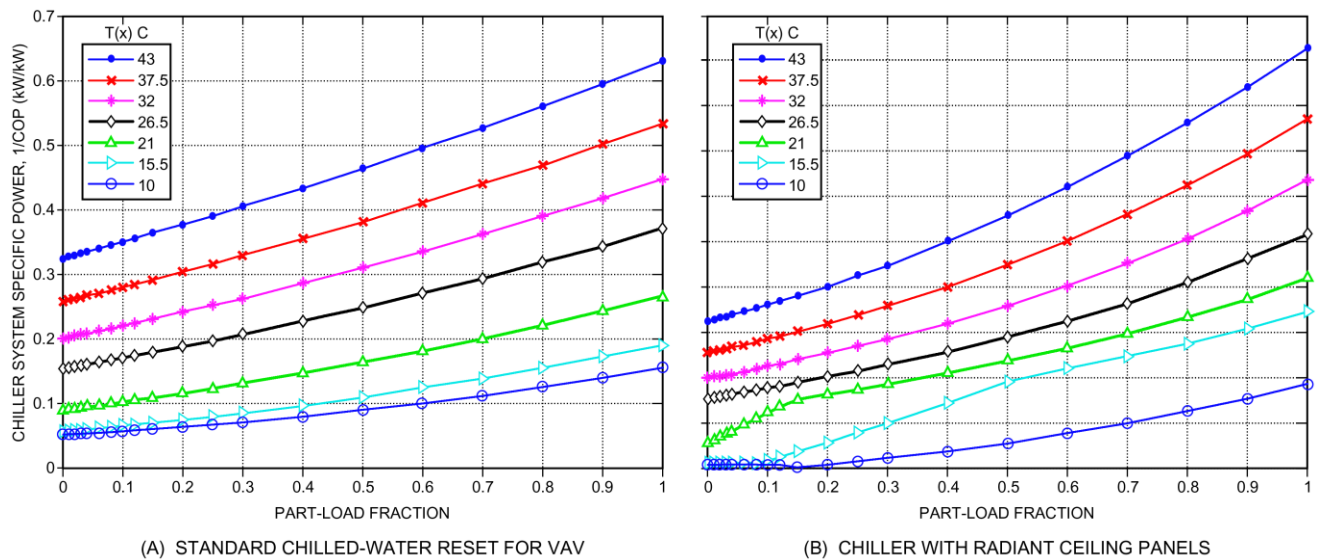
(A) STANDARD CHILLED-WATER RESET FOR VAV   (B) CHILLER WITH RADIANT CEILING PANELS

Fig. 22   Performance of an Optimally Controlled Chiller for Two Different Load-Side Boundary Conditions

(A) NO THERMAL ENERGY STORAGE (TES)   (B) OPTIMAL PRECOOLING OF TES

Fig. 23   Chiller Load Distributions for Chicago

DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*

(A) NO THERMAL ENERGY STORAGE (TES)

(B) OPTIMAL PRECOOLING OF TES
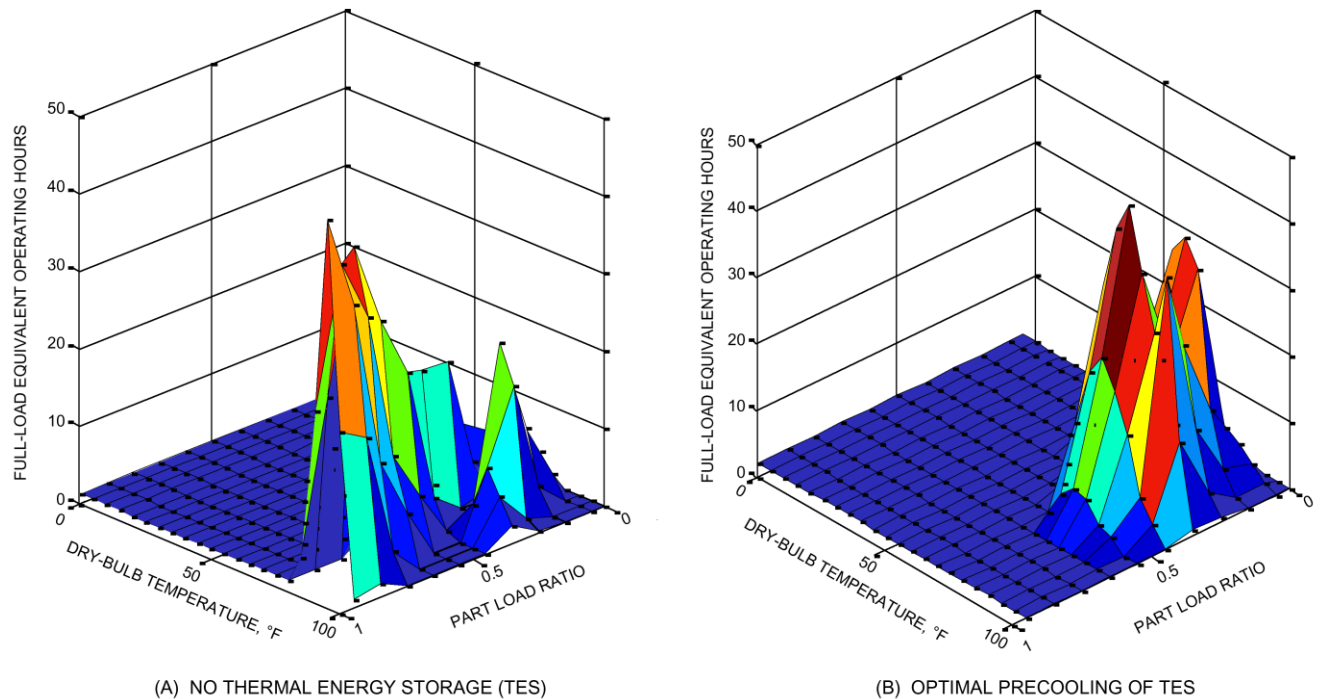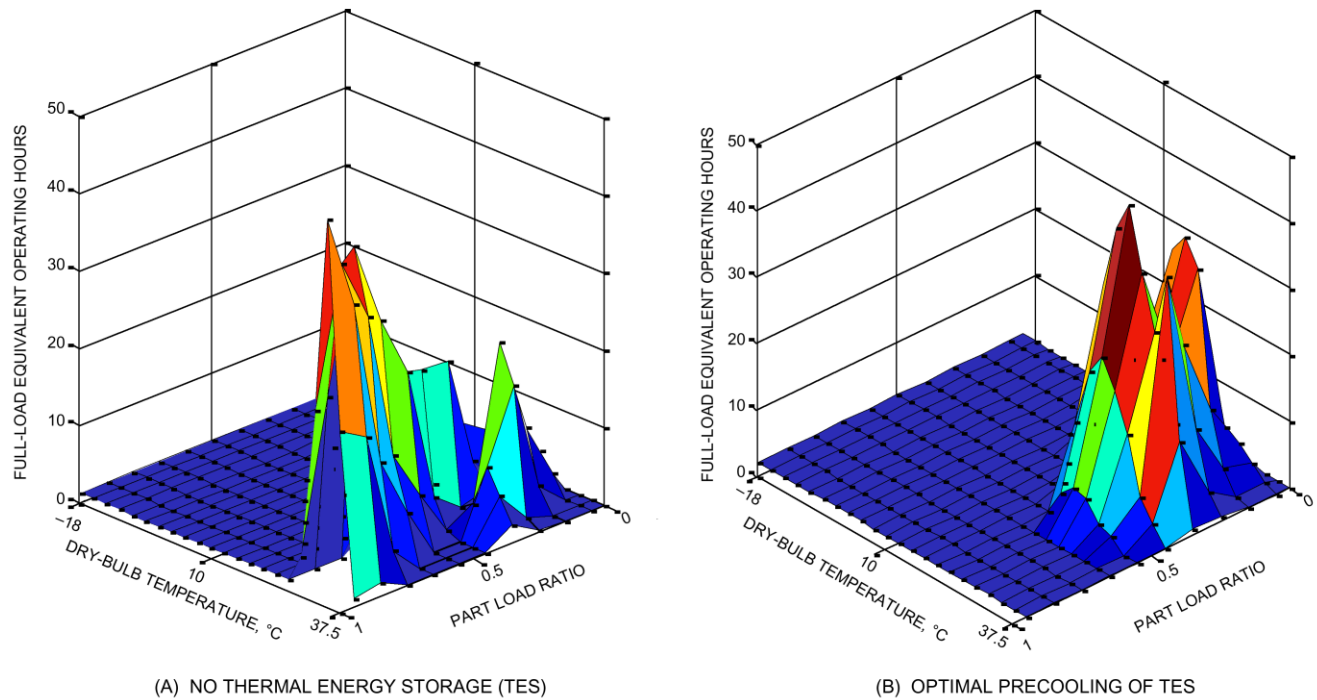
Fig. 23 Chiller Load Distributions for Chicago

The base case for savings used a two-speed chiller built from the same components as the statically optimized chiller and VAV or CAV air-handling unit, depending on building type, with air-side economizer for distribution and no storage. The controller objective function for the TABS storage case, Equation (12), involves only the chiller power used to meet sensible load, $P = u/COP(u,t)$ with $E = 1$ and $D = 0$ to produce the highest energy savings. The need for the upper-bound constraint of Equation 16 was eliminated by providing sufficient mass to satisfy the peak day sensible cooling load and assuming zero charge carryover from day to day, In the simulations, perfect 24 h forecasts of load and outdoor temperature were used; in practice, a model-based predictive control using the forecast methods from this chapter and publicly or commercially available forecasts of weather would result in some loss of performance (Krarti et al. 2007). The results in Table 3 indicate substantial energy savings potential for most of the building types in all 16 climates. For high-performance buildings, the savings were substantial in most building types and climates where economizer cooling potential is modest and the annual sensible load is dominant. Capital costs of the base and low-lift TABS systems were estimated and compared. For large office buildings, the TABS cooling system capital cost was estimated to be less than that of the VAV system.

Table 3 Energy Savings Potential for Precooling with High Part-Load Efficiency Chiller (Katipamula, et al, 2010a)

Building Type	Standard Building			High-Performance Building		
	Minimum	Maximum	Average	Minimum	Maximum	Average
Office, small	59%	77%	70%	−19%	43%	25%
medium	13%	52%	37%	7%	50%	25%
large	21%	61%	40%	−4%	36%	13%
Retail, standalone	56%	73%	66%	31%	50%	41%
strip mall	45%	63%	56%	−7%	41%	16%
Primary school	46%	55%	51%	22%	46%	34%
Secondary school	32%	49%	43%	10%	37%	26%
Hotel, large	16%	57%	44%	−11%	47%	29%
Supermarket	59%	78%	68%	35%	63%	51%
Warehouse	50%	81%	69%	−5%	69%	40%
Outpatient	65%	83%	78%	34%	67%	53%
Hospital	48%	76%	64%	10%	49%	37%

These results are based on simulations in which the control for charging of the TABS is assumed to be modeled exactly. One problem with TABS is that precooling can no longer be reliably controlled by room temperature set-point adjustments. A further problem is that the state of charge cannot be readily measured. For these reasons, supervisory control of a real TABS implementation is considered to require some form of model-based predictive control.

A full-scale laboratory test of TABS cooling with model-based predictive control was conducted by Gayeski (2010b). The test results showed energy savings of 25 to 30% compared to an SEER-16 all-air system for typical Atlanta and Phoenix summer conditions. In each two-week-long test, the same variable-speed compressor/condenser unit was connected to a conventional indoor unit for the baseline case and to a hydronic evaporator supplying chilled water to a 6 in. (150 mm) concrete slab for the optimal precooling case. The results were comparable to the simulation results of Katipamula et al. (2010a) for a medium office building in Atlanta and Phoenix.

Dynamic Optimization of Combined Thermal Energy Storage Systems

Combination of TABS and Direct Sensible Cooling. Zakula (2014a) used the two weeks of transient data from Gayeski's carefully instrumented test room to verify a TRNSYS TABS model linked to a model predictive controller using an equivalent CRTF model (Seem) to represent TABS and zone. With this capability, the seasonal (May 1 to Sep 30) performance of the test room was simulated, first with cooling delivered by TABS only, by optimal scheduling of VAV precooling, and by their combination. Figure 24 shows the percent energy savings for five cities. The savings using TABS only with respect to conventional VAV (no precooling) is shown in 24a. The savings using TABS on top of optimal VAV precooling is shown in 24b. The existence of an exposed floor mass apparently resulted in passive precooling producing savings of about 70% of the savings achieved with TABS. Thus combing TABS with precooling by sensible-only VRF produces larger savings than TABS alone because the TABS no longer has to carry the full load (Zakula 2014b)
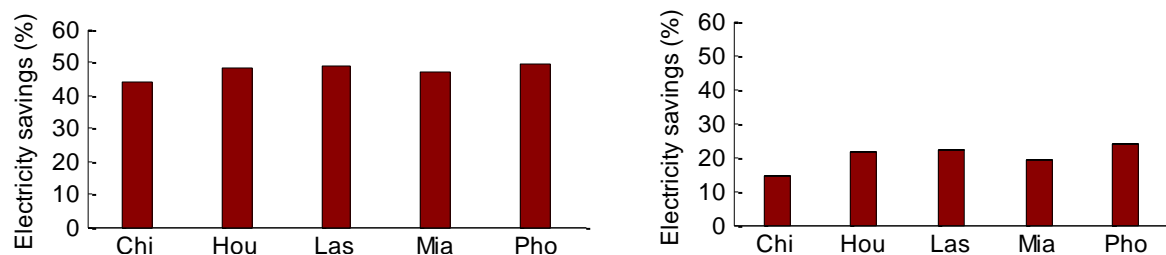

Fig. 23   **Savings With Respect To Conventional VAV (left) and Sensible-Only MPC-VRF (right) for Five Cities**

Niswander (2013) modeled the combination of TABS and sensible-only VRF with model predictive control deciding the sequence of capacity delivered to the TABS and directly to the room over each 24-hour planning horizon. In addition he modeled a third mode in which heat can be rejected from the VRF terminal units to the TABS at very high COP in order to reduce the impact of high RTP rates in mid-late afternoon. The controller objective function, Equation (12), now includes three modes of heat pumpr operation, $P = u_p/\text{COP}(u_p,t) + u_p/\text{COP}(u_p,t) + u_p/\text{COP}(u_p,t)$ where subscript p denotes precooling of TABS, q denotes direct cooling (possibly to effect passive precooling) and r denotes rejection of heat to TABS. Simularions were carried out with $E = 1$ and $D = 0$ to produce the highest energy savings and with RTP rates to produce the largest shifting of load. The upper-bound constraint of Equation 16 was again eliminated but in this analysis the states of both passive storage and TABS were carried from day to day. Some typical peak-shifting results are shown in Figure 25 for flat rate (left) and a mode; of the Com Ed RTP rate (right).
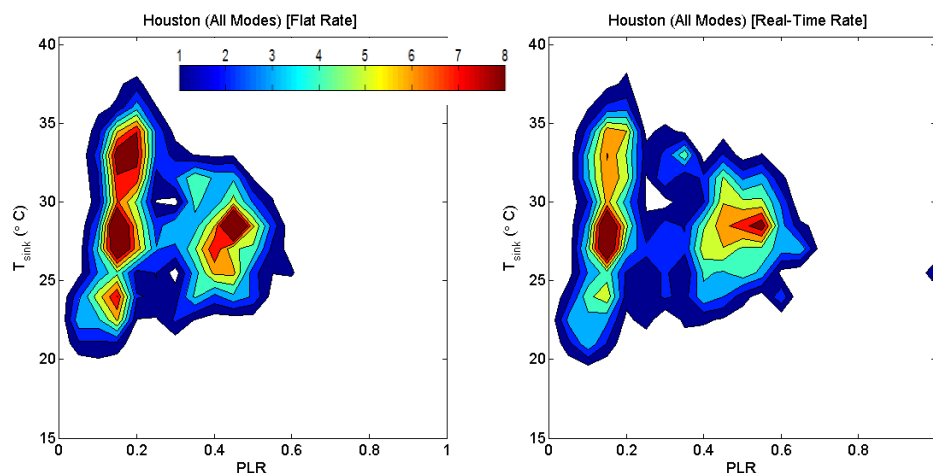
Fig. 23   **FLEOH Distributions with TABS Acting Both as Cool Storage and Demand Responsive Heat Sink**

Combined Passive and Discrete Thermal Storage. Most investigations into the optimal control of combined discrete and passive building thermal storage inventory rely on a detailed white-box or gray-box model of building thermal response and equipment performance [see, e.g., Henze et al. 2005]. However, Liu and Henze (2006a, 2006b) describe a novel approach to optimally control commercial building passive and discrete thermal storage inventory simultaneously: a hybrid control scheme that combines features of model-based optimal control and model-free reinforcement learning control.

Theoretically, the reinforcement learning algorithm, based on Watkins and Dayan (1992), approximates dynamic-programming-based optimal control by sampling the cost space and can reach the true optimum, given properly selected learning parameters and long enough learning time. The amount of required training is not yet realistic if the controller is directly implemented in a commercial building application. This constitutes the major drawback of the reinforcement learning control approach; contextual information in some form needs to be introduced to expedite learning of the fundamental features of the problem, while reinforcement learning fine-tunes the controller. This realization inspired the development of the hybrid learning control scheme (Liu and Henze 2006a).

Liu and Henze (2006b) analyzed the performance of this hybrid controller installed in a full-scale laboratory facility. Operating cost savings from the application of both discrete and passive TES were attained with the hybrid control approach compared with conventional building control; however, the savings were lower than for the case of model-based predictive optimal control. As for the case of model-based predictive control, the hybrid controller's performance is largely affected by the quality of the training model, and extensive real-time learning is required for the learning controller to eliminate any false cues it receives during the initial training period. Nevertheless, compared with standard reinforcement learning, Liu and Henze's proposed hybrid controller is much more readily implemented in a commercial building.

FORECASTING DIURNAL COOLING AND WHOLE-BUILDING DEMAND PROFILES

As discussed previously, forecasts of cooling requirements and electrical use in buildings are often necessary for the control of thermal storage to shift electrical use from on-peak to off-peak periods. In addition, forecasts can help plant operators anticipate major changes in operating modes, such as bringing additional chillers online.

In most methods, predictions are estimated as a function of time-varying input variables that affect cooling requirements and electrical use. Examples of inputs that affect building energy use include (1) ambient dry-bulb temperature, (2) ambient wet-bulb temperature, (3) solar radiation, (4) building occupancy, and (5) wind speed. Methods that include time-varying measured input variables are often termed **deterministic methods**.

Not all inputs affecting cooling requirements and electric use are easily measured. For instance, building occupancy is difficult to determine and solar radiation measurements are expensive. In addition, the accuracy of forecast models that use deterministic inputs depends on the accuracy of future predictions of the inputs. As a result, most inputs that affect building energy use are typically not used.

Much of the time-dependent variation in cooling loads and electrical use for a building can be captured with time as a deterministic input. For instance, building occupancy follows a regular schedule that depends on time of day and time of year. In addition, variations in ambient conditions follow a regular daily and seasonal pattern. Many forecasting methods use time in place of unmeasured deterministic inputs in a functional form that captures the average time dependence of the variation in energy use.

A deterministic model has limited accuracy for forecasts because of both unmeasured and unpredictable (random) input variables. Short-term forecasts can be improved significantly by adding previous values of deterministic inputs and previous output measurements (cooling requirements or electrical use) as inputs to the forecasting model. The time history of these inputs provides valuable information about recent trends in the time variation of the forecasted variable and the unmeasured input variables that affect it. Most forecasting methods use historical variables to predict the future.

Any forecasting method requires that a functional form is defined and the parameters of the model are learned based on measured data. Either offline or online methods can be used to estimate parameters. Offline methods involve estimating parameters from a batch of collected data. Typically, parameters are determined by minimizing the sum of squares of the forecast errors. The parameters of the process are assumed to be constant over time in the offline methods. Online methods allow the parameters of the forecasting model to vary slowly with time. Again, the sum of squares of the forecast errors is minimized, but sequentially or recursively. Often, a forgetting factor is used to give additional weight to the recent data. The ability to track time-varying systems can be important when forecasting cooling requirements or electricity use in buildings, because of the influence of seasonal variations in weather.

Forrester and Wepfer (1984) presented a forecasting algorithm that uses current and previous ambient temperatures and previous loads to predict future requirements. Trends on an hourly time scale are accounted for with measured inputs for a few hours preceding the current time. Day-to-day trends are considered by using the value of the load that occurred 24 h earlier as an input. One of the major limitations of this model is its inability to accurately predict loads when an occupied day (e.g., Monday) follows an unoccupied (e.g., Sunday) or when an unoccupied day follows an occupied day (e.g., Saturday). The cooling load for a particular hour of the day on a Monday depends very little on the requirement 24 h earlier on Sunday. Forrester and Wepfer (1984) described a number of methods for eliminating this 24 h indicator. MacArthur et al. (1989) also presented a load profile prediction algorithm that uses a 24 h regressor.

Armstrong et al. (1989) presented a very simple method for forecasting either cooling or electrical requirements that does not use the 24 h regressor; Seem and Braun (1991) further developed and validated this method. The "average" time-of-day and time-of-week trends are modeled using a lookup table with time and type of day (e.g., occupied versus unoccupied) as the deterministic input variables. Entries in the table are updated using an exponentially weighted, moving-average model. Short-term trends are modeled using previous hourly measurements of cooling requirements in an autoregressive (AR) model. Model parameters adapt to slow changes in system characteristics. The combination of updating the table and modifying model parameters works well in adapting the forecasting algorithm to changes in season and occupancy schedule.

Kreider and Wang (1991) used artificial neural networks (ANNs) to predict energy consumption of various HVAC equipment in a commercial building. Data inputs to the ANN included (1) previous hour's electrical power consumption, (2) building occupancy, (3) wind speed, (4) ambient relative humidity, (5) ambient dry-bulb temperature, (6) previous hour's ambient dry-bulb temperature, (7) two hour's previous ambient temperature, and (8) sine and cosine of the hour number to roughly represent the diurnal change of temperature and solar insolation. The primary purpose in developing these models was to detect changes in equipment and system performance for monitoring purposes. However, the authors suggested that an ANN-based predictor might be valuable when used to predict energy consumption with a network based on recent historical data. Forecasts of all deterministic input variables are necessary to apply this method.

Gibson and Kraft (1993) used an ANN to predict building electrical consumption as part of the operation and control of a thermal energy storage (TES) cooling system. The ANN used the following inputs: (1) electric demand of occupants (lighting and other loads), (2) electric demand of TES cooling tower fans, (3) outdoor ambient temperature, (4) outdoor ambient temperature/inside target temperature, (5) outdoor ambient relative humidity, (6) on/off status for building cooling, (7) cooling system on/off status, (8) chiller #1 direct-cooling mode on/off status, (9) chiller #2 direct-cooling mode on/off status, (10) ice storage discharging mode on/off status, (11) ice storage charging mode on/off status, (12) chiller #1 charging mode on/off status, and (13) chiller #2 charging mode on/off status. To use this forecaster, values of each of these inputs must be predicted. Although the authors suggest that average occupancy demand profile be used as an input, they do not state how the other input variables should be forecast.

A Forecasting Algorithm

This section presents an algorithm for forecasting hourly cooling requirements or electrical use in buildings that is based on the method developed by Seem and Braun (1991). At a given hour n, the forecast value is

$$\hat{E}_{(n)} = \hat{X}_{(n)} + \hat{D}_{(h,d)} \qquad (61)$$

where

$\hat{E}_{(n)}$ = forecast cooling load or electrical use for hour n

$\hat{X}_{(n)}$ = stochastic or probabilistic part of forecast for hour n

$\hat{D}_{(h,d)}$ = deterministic part of forecast at hour n associated with hth hour of day and current day type d

The deterministic part of the forecast is simply a lookup table for the forecasted variable in terms of hour of day h and type of day d. Seem and Braun recommend using three distinct day types: unoccupied days, occupied days following unoccupied days, and occupied days following occupied days. The three day types account for differences between the building responses associated with return from night setup and return from weekend setup. The building operator or control engineer must specify the number of day types and a calendar of day types.

Given the hour of day and day type, the deterministic part of the forecast is simply the value stored in that location in the table. Table entries are updated when a new measurement becomes available for that hour and day type. Updates are accomplished through the use of an exponentially weighted, moving-average (EWMA) model as

$$\hat{D}_{(h,d)} = \hat{D}_{(h,d)old} + \lambda[E(n) - \hat{D}_{(h,d)old}] \qquad (62)$$

where

$E(n)$ = measured value of cooling load or electrical use for current hour n

λ = exponential smoothing constant, $0 < \lambda < 1$

$\hat{D}_{(h,d)old}$ = previous table entry for $\hat{D}_{(h,d)}$

As λ increases, the more recent observations have more influence on the average. As λ approaches zero, the table entry approaches the average of all data for that hour and day type. When λ equals one, the table entry is updated with the most recent measured value. Seem and Braun recommend using a value of 0.30 for λ in conjunction with three day types and 0.18 with two day types.

The stochastic portion of the forecast is estimated with a third-order autoregressive model, AR(3), of forecasting errors associated with the deterministic model. With this model, an estimate of the next hour's error in the deterministic model forecast is given by the following equation:

$$\hat{X}_{(n+1)} = \phi_1 X(n) + \phi_2 X(n-1) + \phi_3 X(n-2) \qquad (63)$$

where

$X(n)$ = difference between measurement and deterministic forecast of cooling load or electrical use at any hour n

ϕ_1, ϕ_2, ϕ_3 = parameters of AR(3) model that must be learned

The error in the deterministic forecast at any hour is simply

$$X(n) = E(n) - \hat{D}_{(h,d)} \qquad (64)$$

For forecasting more than one hour ahead, conditional expectation is used to estimate the deterministic model forecast errors using the AR(3) model as follows:

$$\hat{X}(n+2) = \phi_1 \hat{X}(n+1) + \phi_2 X(n) + \phi_3 X(n-1)$$

$$\hat{X}(n+3) = \phi_1 \hat{X}(n+2) + \phi_2 \hat{X}(n+1) + \phi_3 X(n)$$

$$\vdots$$

$$\hat{X}(n+k) = \phi_1 \hat{X}(n+k+1) + \phi_2 \hat{X}(n+k-2)$$
$$+ \phi_3 \hat{X}(n+k-3) \qquad \text{for } k > 3$$

$$(65)$$

Online estimation of the AR(3) model parameters is accomplished by minimizing the following time-dependent cost function:

$$J(\phi) = \sum_{k=1}^{n} \alpha^{n-k}[X(k) - \hat{X}(k)]^2$$

(66)

where the constant α is called the forgetting factor and has a value between 0 and 1. With this formulation, the residual for the current time step has a weight of one and the residual for k time steps back has a weight of α_k. By choosing a value of α that is positive and less than one, recent data have greater influence on the parameter estimates. In this manner, the model can track changes due to seasonal or other effects. Seem and Braun recommend using a forgetting factor of 0.99. Parameters of the AR(3) model should be updated at each hour when a new measurement becomes available.

Ljung and Söderström (1983) describe online estimation methods for determining coefficients of an AR model. The parameter estimates should be evaluated for stability. If an AR model is not stable, then the forecasts will grow without bound as the time of forecasts increases. Ljung and Söderström discuss methods for checking stability.

Seem and Braun compared forecasts of electrical usage with both simulated and measured data. Figure 42 shows the standard deviation of the 1 h through 24 h errors in electrical use forecasts for annual simulation results. Results are given for the deterministic model alone, deterministic plus AR(2), and deterministic plus AR(3). For the combined models, the standard deviation of the residuals increases as the forecast length increases. For short time steps (i.e., less than 6 h), the combined deterministic and stochastic models provide much better forecasts than the purely deterministic model (i.e., lookup table).

Seem and Braun also investigated a method for adjusting the deterministic forecast based on using weather service forecasts of maximum daily ambient temperature as an input. For short periods (i.e., less than 4 h), forecasts for the temperature-dependent model were nearly identical to forecasts for the temperature-independent model. For longer periods, the temperature-dependent model provided better forecasts than the temperature-independent model.
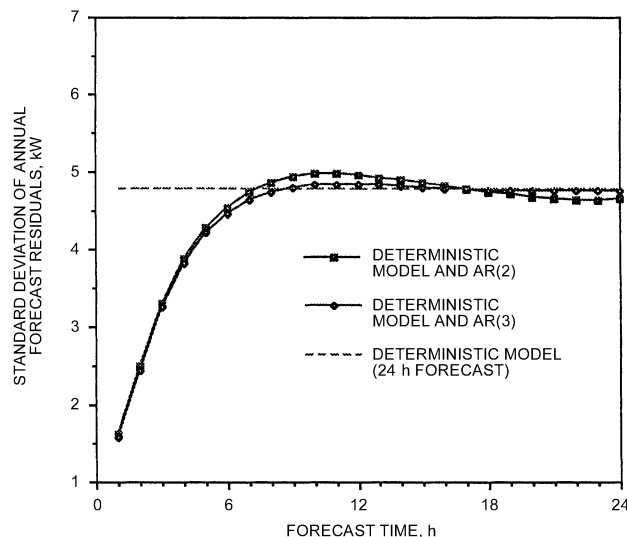

Fig. 42   Standard Deviation of Annual Errors for 1 to 24 h Forecasts

BLACK BOX PREDICTIVE COOLING CONTROL STRATEGIES

Model predictive controls applied to commercial buildings require short-term weather forecasts to optimally adjust set points in a supervisory control environment.

Some researchers are convinced that nonlinear forecasting models based on neural networks (NNs) provide superior performance over traditional time-series analysis. Florita and Henze (2009) sought to identify the complexity required for short-term weather forecasting in the context of a model predictive control environment. Moving average (MA) models with various enhancements and neural network models were used to predict weather variables seasonally in numerous geographic locations. Their performance was statistically assessed using the coefficient of variation (CV) and mean bias error (MBE) values. When a cyclical two-stage model predictive control process of policy planning followed by execution was used, Florita and Henze found that even the most complicated nonlinear autoregressive neural network with exogenous input does not appear to warrant the additional efforts in forecasting model development and training in comparison to the simpler MA models.

Predictive control strategies for both passive and discrete thermal storage systems have been field-tested (Krarti et al. 2007; Morgan and Krarti 2010). Morgan and Krarti (2010) investigated performance of various control strategies for combined passive and dicrete TES systems in a Colorado elementary school equipped with an ice storage system. The predictive optimal control strategies were developed using an EnergyPlus-based simulation environment (Krarti et al. 2007; Zhou et al. 2005). The simulation environment was found to be effective in defining and implementing predictive optimal controls for both passive and discrete TES systems for the buildings. Figures 43 and 44 show examples of the field testing results for the building energy use performed by Morgan and Krarti (2010).
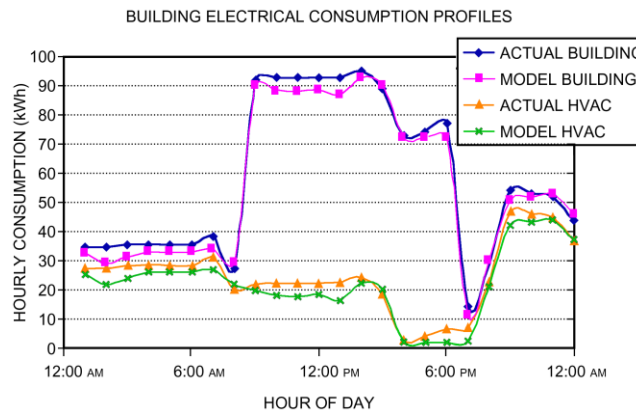
Fig. 43 Building Electricity Use Profiles for 6 h Predictive Optimal Control
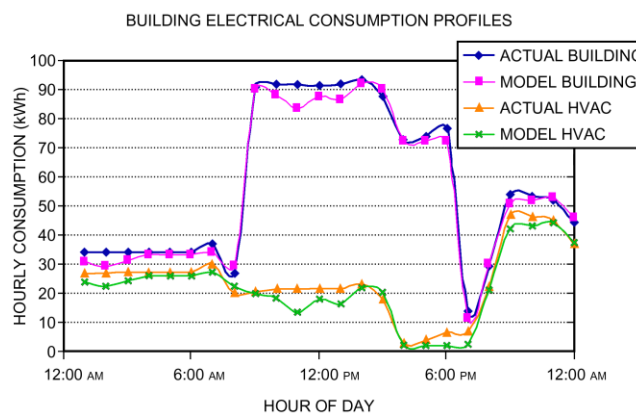

Fig. 44 Building Electricity Use Profiles for 24 h Predictive Optimal Control

CONTROL STRATEGIES FOR HEATING SYSTEMS

Boiler efficiency depends on many factors, such as combustion airflow rate, load factor, and water temperature in hot-water boilers (or pressure for steam boilers). Opportunities for energy and cost reduction in boiler plants by supervisory control include sequencing and loading of multiple boilers and resetting the hot-water supply temperature set point (for hot-water boilers) or the steam pressure set point (for steam boilers) (Dyer and Maples 1981).

Sequencing and Loading of Multiple Boilers

Generally, boilers operate most efficiently at a 65 to 85% full-load rating. Boiler efficiencies fall off at higher and lower load points, with the decrease most pronounced at low load conditions. Boiler efficiency can be calculated by means of stack temperature and percent O_2 (or percent excess air) in the boiler stack for a given fuel type. Part-load curves of boiler efficiency versus hot-water or steam load should be developed for each boiler. These curves could be dynamically updated at discrete load levels based on the hot-water or steam plant characteristics to allow the control strategy to continuously predict the input fuel requirement for any given heat load. When the hot-water temperature or steam pressure drops below set point for the predetermined time interval (e.g., 5 min), the most efficient combination of boilers must be selected and turned on to meet the load. The least efficient boiler should be shut down and banked in hot standby if its capacity drops below the spare capacity of the current number of boilers operating (or, for primary/secondary hot-water systems, if the flow rate of the associated primary hot-water pump is less than the difference between primary and secondary hot-water flow rates) for a predetermined time interval (e.g., 5 min). The spare capacity of the current online boilers is equal to their full-load capacity minus the current hot water load.

Load Conditions for Bringing Boilers Online or Offline

The specifics of the strategy for bringing boilers online depend on the type of boiler. Hot-water boilers have dedicated or nondedicated hot-water pumps; steam boilers do not have hot-water pumps, but rely on differences in steam pressure between the boiler steam header discharge and the point of use to distribute steam throughout the system.

Hot-Water Boilers with Dedicated Pumps. The strategy for hot-water boilers with dedicated hot-water pumps is similar to that for chillers with dedicated chilled-water and condenser water pumps: another boiler should be brought online when operating boilers reach

capacity, because the efficiency of the boiler should include the power to drive its associated hot-water pump. This can be determined when hot-water temperature drops below its set point for a predetermined time interval (e.g., 5 min).

Hot-water boilers with dedicated pumps can be brought online and offline with the following logic:

1. Continuously calculate the load ratio of each boiler or boiler combination. For the ith boiler,

$$\text{LR}_i = \frac{\dfrac{Q_{load}}{Q_{blr,des,i}}}{} \tag{44}$$

where LR_i is load ratio of the ith boiler combination, Q_{load} is the total boiler plant load, and $Q_{blr,des,i}$ is the design (rated) output of the ith boiler combination.

2. Every sampling interval (e.g., 60 s), calculate the predicted input fuel requirement for each boiler combination as

$$\text{IF}_i = \frac{(\text{LR}_i)\dot{Q}_{blr,des,i}}{\eta_i} \tag{45}$$

where IF_i is the input fuel requirement and η_i is the efficiency of the ith boiler combination.

3. Continuously evaluate time-averaged values of the hot-water supply temperature over a fixed time interval (e.g., 5 min).

4. If the hot-water supply temperature drops below its set point for a predetermined time interval (e.g., 5 min), then, from boiler part-load performance curves, select the boiler combination with a load ratio between 0.5 and 1.0 and with the least input fuel requirement to meet the load, and turn this combination of boilers on. Note that this strategy greatly reduces the possibility of short-cycling boilers because the new combination of boilers to be started likely includes boilers already operating (i.e., only one additional boiler is likely to be added).

5. If the capacity of the least efficient online boiler drops below the spare capacity of the current number of boilers operating (or for a primary/secondary hot-water system, if the flow rate of the associated primary hot-water pump is less than the difference between primary and secondary hot-water flow rates) for a predetermined time interval (e.g., 5 min), then shut down and bank this boiler in hot standby.

Hot-Water Boilers with Nondedicated Pumps or Steam Boilers. For hot-water systems without dedicated hot-water pumps or for steam systems, the optimal load conditions for bringing boilers online or offline do not generally occur at the full capacity of the online boilers. For these systems, a new boiler combination should be brought online whenever the hot-water supply temperature or steam pressure falls below set point for a predetermined time interval (e.g., 5 min) *and* the part-load efficiency curves of the boiler combination predicts that the new combination of boilers can meet the required load using significantly less (e.g., 5%) input fuel.

Optimal Boiler Load Distribution

Optimal load distribution strategies for boilers are similar to those for chillers. For boilers with similar performance characteristics, the optimal boiler loading is similar to Equation (35) for chillers:

$$\dot{Q}^*_{blr,i} = \frac{\dot{Q}_{load}}{\displaystyle\sum_{i=1}^{N} Q_{blr,des,i}} \dot{Q}_{blr,des,i} \tag{46}$$

where $\dot{Q}^*_{blr,i}$ is the optimal load for the ith boiler.

For boilers with significantly different performance characteristics, the criterion for optimal boiler loading is similar to Equations (36) and (37) for chillers, except that boiler cost of operation is used:

$$\frac{\partial C_{blr,i}}{\partial \dot{Q}_{blr,i}} = \frac{\partial C_{blr,j}}{\partial \dot{Q}_{blr,j}} \quad \text{for all } i \text{ and } j \tag{47}$$

and subject to the constraint that

$$\sum_{i=1}^{N} \dot{Q}_{blr,i} = \dot{Q}_{load} \tag{48}$$

where $C^{blr,i}$ and $C_{blr,j}$ are the operating costs of boiler i and j, respectively, $\dot{Q}_{blr,i}$ is heating load for the ith boiler, and $\dot{Q}_{load}$ is total heating load.

In general, the boiler operating-cost curve can be calculated as a quadratic function of heating load only. For the ith boiler,

$$C_{blr,i} = b_{0,i}\dot{Q}^2_{blr,i} + b_{1,i}\dot{Q}_{blr,i} + b_{2,i} \tag{49}$$

Applying the criterion of Equation (47) to Equation (49),

$$2b_{0,i}\dot{Q}^*_{blr,i} + b_{1,i} = 2b_{0,j}\dot{Q}^*_{blr,\bullet j} + b_{1,j} \tag{50}$$

where $\dot{Q}^*_{blr,i}$ is the optimal load for the ith boiler.

Maintaining Boilers in Standby Mode

It is generally more economical to run fewer boilers at a high rating. However, the integrity of the steam or hot-water supply must be maintained in the event of a forced outage of one of the operating boilers or if the facility experiences highly diverse load swings throughout the heating season. Both conditions can often be satisfied by maintaining a boiler in standby or "live bank" mode. For example, in this mode, a steam boiler is isolated from the steam system at no load but is kept at system operating pressure by periodic firing of either the igniters or a main burner to counteract ambient heat losses.

Supply Water and Supply Pressure Reset for Boilers

Standby losses are reduced and overall efficiencies enhanced by operating hot-water boilers at the lowest acceptable temperature. Condensing boilers achieve significantly higher combustion efficiencies at water temperatures below the flue gas dew point when they are operating in condensing mode (see Chapter 31 in the 2008 *ASHRAE Handbook—HVAC Systems and Equipment*). Hot-water boilers of this type are very efficient at part-load operation when a high water temperature is not required. Energy savings are therefore possible if the supply water temperature is maintained at the minimum level required to satisfy the largest heating load. However, to minimize condensation of flue gases and consequent boiler damage from acid, water temperature should not be reset below that recommended by the boiler manufacturer (typically 140°F [60°C]). Similarly, energy can be saved in steam heating systems by maintaining supply pressure at the minimum level required to satisfy the largest heating load.

Simple control strategies can be used to generate a suboptimal hot-water temperature (for hot-water boilers) or steam pressure (for steam boilers). An energy management and control system must be interfaced to the boiler controls and be capable of monitoring the position of the valve controlling the flow of hot water to the heating coils or steam pressure at the most critical zone. For a hot-water system,

1. Continuously monitor the hot-water valve position of the various heating zones.
2. If none of the hot-water valves are greater than 95% open, lower boiler hot-water supply temperature by a small increment (e.g., 1°F [0.5 K]) each reset time interval (a predetermined interval established by system thermal lag characteristics, e.g., 15 to 20 min).
3. Once one hot-water valve opens beyond 95%, stop downward resets of boiler hot-water temperature set point.
4. If two or more hot-water valves open beyond 95%, raise boiler hot-water temperature set point by a small increment (e.g., 1°F [0.5 K]) each reset interval.

For a steam system, the steam header pressure should be lowered to a value that just satisfies the highest pressure demand. *Caution*: for nongravity condensate return systems, steam pressure reset could impede condensate return (see Chapter 10 of the 2008 *ASHRAE Handbook—HVAC Systems and Equipment*).

Operating Constraints

Note that there are practical limitations on the extent of automatic operation if damage to the boiler is to be prevented. Control strategies to reduce boiler energy consumption can also conflict with recommended boiler operating practice. For example, in addition to the flue gas condensation concerns mentioned previously, rapid changes in boiler jacket temperature (thermal shock) brought about by abrupt changes in boiler water temperature or flow, firing rate, or air temperature entering the boiler should be avoided. The repeated occurrence of such transient conditions may weaken the metal and lead to cracking and/or loose tubes. It is therefore important to follow all of the recommendations of the American Boiler Manufacturers Association (ABMA 1998).

CONTROL STRATEGIES FOR AIR-HANDLING UNITS

Air Handler Sequencing and Economizer Cooling

Traditional air handler sequencing strategies use a single PI controller to control heating, cooling with outdoor air, mechanical cooling with 100% outdoor air, and mechanical cooling with minimum outdoor air. Sequencing between these different modes is accomplished by splitting the controller output into different regions of operation, as shown in Figure 37.
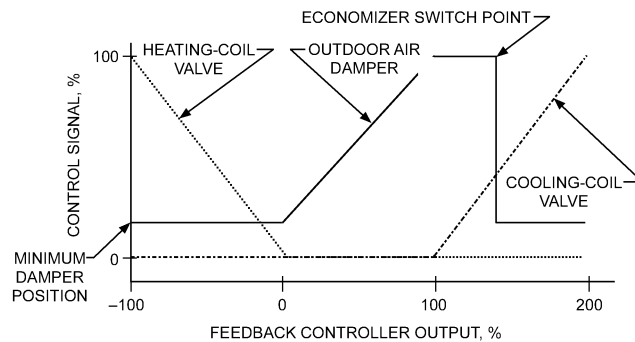
Fig. 37 AHU Sequencing Strategy with Single Feedback Controller

Figure 37 depicts the relationship between the control signal to the valves and dampers and the feedback controller output. The controller adjusts its output to maintain the supply air temperature set point. If output is between 100 and 200%, mechanical cooling is used. When outdoor conditions are suitable, the outdoor air dampers switch from minimum position (minimum ventilation air) to fully open. For a dry-bulb economizer, this switch point occurs when ambient air is less than a specified value. This switch point should be less than the switch point to return to minimum outdoor air, to ensure stable control. The economizer switchover temperature may be significantly lower than the return air temperature (e.g., 10°F [5 K] lower) in humid climates where latent ventilation loads are significant. However, in dry climates, the switchover temperature may be close to the return temperature (e.g., 75°F [24°C]). An enthalpy (or wet-bulb) economizer compares outdoor and return air enthalpies (or wet-bulb temperatures) to initiate or terminate economizer operation. In general, enthalpy economizers yield lower energy costs than dry-bulb economizers, but require measurements of outdoor and return air humidity. Humidity sensors require regular maintenance to ensure accurate readings. When controller output is between 0 and 100% (see Figure 37), the cooling coil valve is fully closed and cooling is provided by ambient air only. In this case, the controller output modulates the position of the outdoor air dampers to maintain the set point. If the controller output signal is between −100 and 0%, the heating coil is used to maintain set point and the outdoor air dampers are set at their minimum position.

A single feedback controller is difficult to tune to perform well for all four modes of operation associated with an AHU. An alternative to the traditional sequencing strategy is to use three separate feedback controllers, as described by Seem et al. (1999). This approach can improve temperature control, reduce actuator usage, and reduce energy costs. Figure 38 shows a state transition diagram for implementing a sequencing strategy that incorporates separate feedback controllers.

In state 1, a feedback controller adjusts the heating valve to maintain the supply air set-point temperature with minimum outdoor air. The transition to state 2 occurs after the control signal has been saturated at the no-heating position for a period equal to a specified state transition delay (e.g., 3 min). In state 2, a second feedback controller adjusts the outdoor and return air dampers to achieve set point with heating and cooling valves closed. Transition back to state 1 only occurs after the damper control signal is saturated at its minimum value for the state transition delay, whereas transition to state 3 is associated with saturation at the maximum damper position for the state transition delay. In state 3, the outdoor air damper remains fully open and a third feedback controller is used to adjust the flow of cooling water to maintain the supply air temperature at set point. Transition back to state 2 occurs if the controller output is saturated at its minimum value for the state transition delay. For a dry-bulb economizer, transition to state 4 occurs when the ambient dry-bulb temperature is greater than the switchover temperature by a dead band (e.g., 2°F [1 K]). The feedback controller continues to modulate the cooling-coil valve to achieve set point. Transition back to state 3 occurs when the ambient dry-bulb is less than the switchover temperature (e.g., 65°F [18°C]). For an enthalpy economizer, the ambient enthalpy is compared with return air enthalpy to initiate transitions between states 3 and 4.
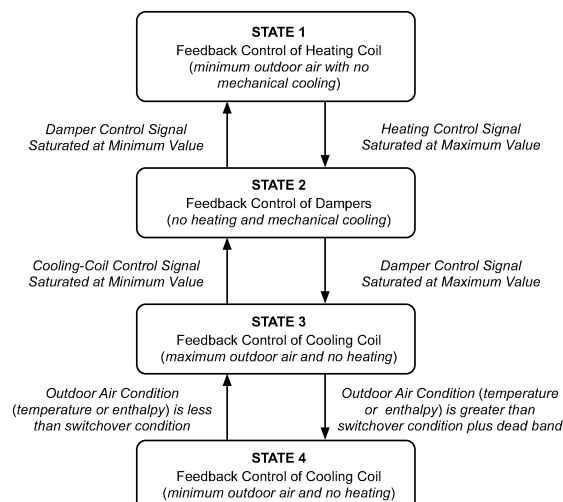
Fig. 38   AHU Sequencing Strategy with Multiple Feedback Controllers

Supply Air Temperature Reset for Constant Air Volume (CAV)

The benefits of resetting supply air temperature set points for CAV systems are significant. Increasing the supply air set point for cooling reduces both the cooling and reheat required, but does not change fan energy. In general, the set point for a CAV system could be set at the highest value that will keep all zone temperatures at their set points and all humidities within acceptable limits. A simple reset strategy based on this concept follows.

At each the decision interval (e.g., 5 min), the following logic can be applied:

1. Check controller outputs for representative zone reheat units and determine time-averaged values over the last decision interval.
2. If any controller output is less than a threshold value (e.g., 5%), then decrease the supply air set point by a fixed value (e.g., 0.5°F [0.25 K]) and go to step 4. Otherwise, go to step 3.
3. If all zone humidities are acceptable and all controller outputs are greater than a threshold value (e.g., 10%), then increase the discharge air set point by a fixed value (e.g., 0.5°F [0.25 K]) and go to step 4. Otherwise, do not change the set point.
4. Limit the set point between upper and lower limits based on comfort considerations.

Static Pressure Reset for Variable Air Volume (VAV)

Flow may be modulated in a VAV system by using dampers on the outlet side of the fan, inlet vanes on the fan, vane-axial fans with controllable pitch fan blades, or variable-speed control of the fan motor. Typically, inputs to any of these controlled devices are modulated to maintain a duct static pressure set point as described in Chapter 47. In a single-duct VAV system, the duct static pressure set point is typically selected by the designer to be a fixed value. The sensor is located at a point in the ductwork such that the established set point will ensure proper operation of the zone VAV boxes under varying load (supply airflow) conditions. A shortcoming of this approach is that static pressure is controlled based on a single sensor intended to represent the pressure available to all VAV boxes. Poor location or malfunction of this sensor will cause operating problems.

For a fixed static pressure set point, all of the VAV boxes tend to close as zone loads and flow requirements decrease. Therefore, flow resistance increases with decreasing load. Significant fan energy savings are possible if the static pressure set point is reset so that at least one of the VAV boxes remains open. With this approach, flow resistance remains relatively constant. Englander and Norford (1992), Hartman (1993), Warren and Norford (1993), and Wei et al. (2004) proposed several different strategies based on this concept. Englander and Norford used simulations to show that either static pressure or fan speed can be controlled directly using a flow error signal from one or more zones and simple rules. Their technique forms the basis of the following reset strategy.

At each decision interval (e.g., 5 min), the following logic can be applied:

1. Check the controller outputs for representative VAV boxes and determine time-averaged values over the last decision interval.
2. If any of the controller outputs are greater than a threshold value (e.g., 98%), then increase the static pressure set point by a fixed value (e.g., 5% of the design range) and go to step 4. Otherwise, go to step 3.
3. If all controller outputs are less than a threshold value (e.g., 90%), then decrease the static pressure set point by a fixed value (e.g., 5% of the design range) and go to step 4. Otherwise, do not change the set point.
4. Limit the set point between upper and lower limits based on upper and lower flow limits and duct design.

CONTROLS STRATEGIES FOR BUILDING ZONES

Recovery from Night Setback or Setup

For buildings that are not continuously occupied, significant savings in operating costs may be realized by raising the building set-point temperature for cooling (setup) and by lowering the set point for heating (setback) during unoccupied times. Bloomfield and Fisk (1977) showed energy savings of 12% for a heavyweight building and 34% for a lightweight building.

An optimal controller for return from night setback or setup returns zone temperatures to the comfort range precisely when the building becomes occupied. Seem et al. (1989b) compared seven different algorithms for minimum return time. Each method requires estimating parameters from measurements of the actual return times from night setback or setup.

Seem et al. (1989b) showed that the optimal return time for cooling was not strongly influenced by the outdoor temperature. The following quadratic function of the initial zone temperature was found to be adequate for estimating the return time:

$$\tau = a_0 + a_1 t_{z,i} + a_2 t_{z,i}^2 \qquad (51)$$

where τ is an estimate of optimal return time, $t_{z,i}$ is initial zone temperature at the beginning of the return period, and a_0, a_1, and a_2 are empirical parameters. The parameters of Equation (51) may be estimated by applying linear least-squares techniques to the difference between the actual return time and the estimates. These parameters may be continuously corrected using recursive updating schemes as outlined by Ljung and Söderström (1983).

For heating, Seem et al. found that ambient temperature has a significant effect on the return time and that the following relationship works well in correlating return times:

$$\tau = a_0 + (1 - w)(a_1 t_{z,i} + a_2 t_{z,i}^2) + w a_3 t_a \qquad (52)$$

where t_a is the ambient temperature, a_0, a_1, a_2, and a_3 are empirical parameters, and w is a weighting function given by

DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*

$$w = 1000^{-(t_{z,i} - t_{unocc})/(t_{occ} - t_{unocc})} \tag{53}$$

where t_{unocc} and t_{occ} are the zone set points for unoccupied and occupied periods. Within the context of Equation (52), this function weights the outdoor temperature more heavily when the initial zone temperature is close to the set-point temperature during the unoccupied time. Again, the parameters of Equation (52) may be estimated by applying linear least-squares techniques to the difference between the actual return time and the estimates.

Ideally, separate equations should be used for zones that have significantly different return times. Equipment operation is initiated for the zone with the earliest return time. In a building with a central cooling system, the equipment should be operated above some minimum load limit. With this constraint, some zones need to be returned to their set points earlier than the optimum time.

Optimal start algorithms often use a measure of the building mass temperature rather than the space temperature to determine return time. Although use of space temperature results in lower energy costs (i.e., shorter return time), the mass temperature may result in better comfort conditions at the time of occupancy.

Emergency Strategy to Limit Peak Cooling Requirements

Keeney and Braun (1997) developed a simple control strategy that uses building thermal mass to reduce peak cooling requirements in the event of a loss of a chiller. This emergency strategy is used only on days where cooling capacity is not sufficient to keep the building in the comfort range using night setup control. It involves precooling the building during unoccupied times and allowing the temperature to float through the comfort zone during occupancy.

The precooling control strategy is depicted in Figure 39 along with conventional night setup control. Precooling is controlled at a constant temperature set point t_{pre}. The warm-up period is used to reset the zone air temperature set point so that the cooling system turns off without calling for heating. During this time, the zone air is warmed by lighting and equipment loads. The occupied set point t_{occ} is set at the low end of the comfort region so that the building mass charge is held as long as cooling capacity is available. This set point is maintained until the limit on cooling capacity is reached. After this point, temperatures in the zones float up and the building thermal mass provides additional cooling. If the precooling and occupied set points have been chosen properly and the cooling capacity is sufficient, zone conditions will remain comfortable throughout the occupied period. The peak cooling requirement can be reduced by as much as 25% using this strategy as compared with night setup control. Thus, the loss of one of four identical chillers could be tolerated. This strategy could also be used for spaces such as auditoriums that have a high occupancy density for a short period.
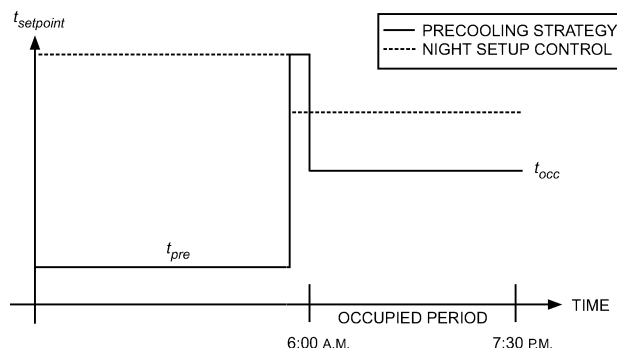

**Fig. 39 Zone Air Temperature Set Points**

The length of time and temperature for precooling and the occupied temperature set point chosen for this strategy strongly influence the capacity reduction and could affect occupant comfort. A reasonable strategy is to precool at 68°F (20°C) beginning at midnight, allow a 30 min warm-up period before occupancy, and then adjust the occupied set point to 70°F (21°C). The zone temperature will then rise above this set point when the chillers are operating at capacity.

Case Study. The control strategy was tested in a 1.4 million square foot (130 000 m²) office building located near Chicago, Illinois. The facility has two identical buildings with very similar internal gains and solar radiation loads, connected by a large, separately cooled entrance area. During tests, the east building used the existing building control strategy and the west building used the precooling strategy.

Four 900 ton (3200 kW) vapor compression chillers normally provide chilled water to the air-handling units. Loss of one chiller results in a 25% reduction of the total capacity. This condition was simulated by limiting the vane position of the two chiller units that cool the west building to 75%. The capacity limitation was imposed directly at the chiller control panels. Set points were provided to local zone controllers from a modern energy management and control system. Chiller cooling loads and zone thermal comfort conditions were monitored throughout the tests.

Consistent with simulation predictions, the precooling control strategy successfully limited the peak load to 75% of the cooling capacity for the west building, whereas the east building operated at 100% of capacity. Figure 40 shows the total chiller coil load for the east and west buildings for a week of testing in the middle of August 1995. The cooling-coil load profile on Monday is the most dramatic example of load shifting during this test period. The peak cooling load for this facility often occurs on Monday morning. The cooling limit was achieved on Monday during a period in which a heat emergency had been declared in the city. The severe ambient conditions were compounded by a power outage that caused a loss of the west-side chiller units for approximately 20 min. Under these demanding

conditions, the precooling strategy maintained occupant comfort while successfully limiting cooling demand of the west side of the building to less than 75% of that for the east side.
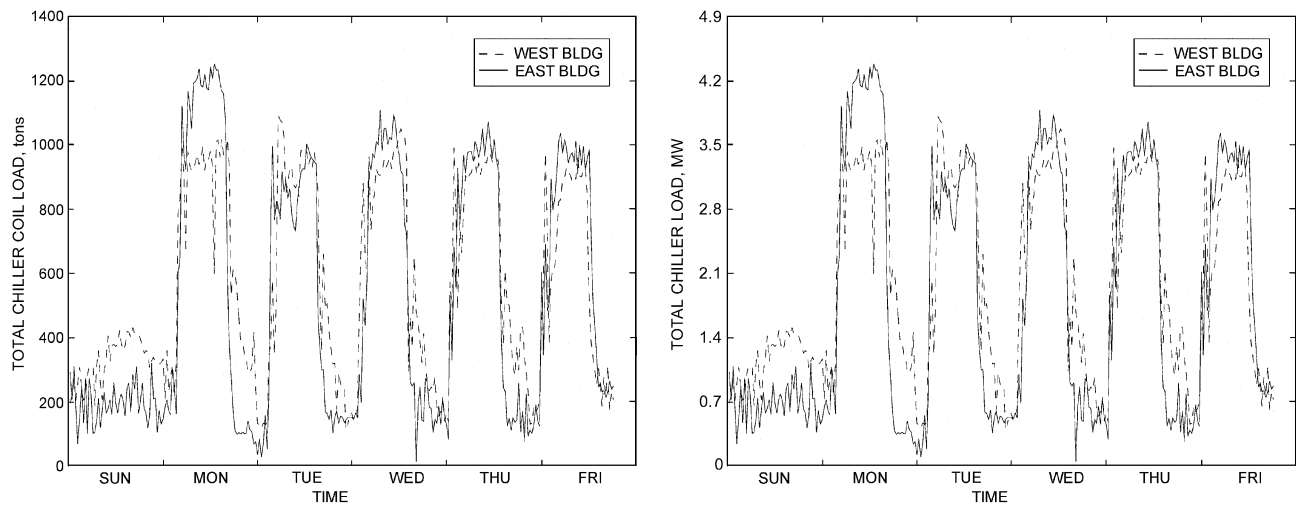
Fig. 40   Total Coil Load for East and West Chiller Units

The east-side cooling requirement was at or below the 75% chiller capacity target for Tuesday through Friday, so the emergency precooling strategy was not necessary. For these off-design days, the emergency strategy was not effective in reducing the on-peak cooling requirements because discharge of the mass was not initiated when capacity was below the target. The thermal mass remained charged so that peak reduction would occur if the target value on the off-design days were reset to a lower value.

Precooling the top floor of the facility had already been implemented into the conventional control strategy used for the east building. This was necessary to maintain comfort conditions with full cooling capacity on hot days. As a result, even greater peak reduction would have been recorded if the precool strategy had been compared with conventional night setup control. The total electrical use was greater for the precooled west building; however, the strategy was designed for an emergency and does not attempt to minimize costs.

This emergency strategy should only be applied on days when the available cooling capacity is not sufficient to maintain comfort conditions when using night setup control. Otherwise, the costs of providing cooling could increase significantly. ASHRAE research project RP-1313 (Cheng et al. 2008; Henze et al. 2007, 2008) developed simplified strategies for optimal control of building zone set points to minimize operating costs.

REFERENCES

ABMA. 1998. *Guideline for the integration of boilers and automated control systems in heating applications*. American Boiler Manufacturers Association, Arlington, VA.

Andresen, I. and M.J. Brandemuehl. 1992. Heat storage in building thermal mass: A parametric study. *ASHRAE Transactions* 98(1).

Armstrong, P.R., T.N. Bechtel, C.E. Hancock, S.E. Jarvis, J.E. Seem, and T.E. Vere. 1989. *Environment for structured implementation of general and advanced HVAC controls—Phase II*. Ch. 7: Small business innovative research program. DOE Contract DE-AC02-85ER 80290.

Armstrong, P.R., S. Katipamula, W. Jiang, D. Winiarski, and L.K. Norford. 2009. Efficient low-lift cooling with radiant distribution, thermal storage and variable-speed chiller controls, Part II: Annual energy use and savings. *HVAC&R Research* 15(2):402-432.

ASHRAE. 2010. Energy standard for buildings except low-rise residential buildings. ANSI/ASHRAE/IESNA *Standard* 90.1-2010.

ASHRAE. 2004. Method of testing the performance of cool storage systems. ANSI/ASHRAE *Standard* 150-2000 (RA 2004).

BEI. 1991. *Boiler efficiency improvement.* Boiler Efficiency Institute, Auburn, AL.

Bellman, R. 1957. *Dynamic programming.* Princeton University Press, Princeton, NJ.

Bloomfield, D.P. and D.J. Fisk. 1977. The optimization of intermittent heating. *Buildings and Environment* 12:43-55.

Brandemuehl, M.J. and J. Bradford. 1998. Implementation of on-line optimal supervisory control of cooling plants without storage (RP-823). ASHRAE Research Project, *Final Report*.

Braun, J.E. 1988. *Methodologies for the design and control of central cooling plants.* Ph.D. dissertation, University of Wisconsin-Madison.

Braun, J.E. 1990. Reducing energy costs and peak electrical demands through optimal control of building thermal storage. *ASHRAE Transactions* 96(1).

Braun, J.E. 1992. A comparison of chiller-priority, storage-priority, and optimal control of an ice-storage system. *ASHRAE Transactions* 98(1): 893-902.

Braun, J.E. 2007a. A near-optimal control strategy for cool storage systems with dynamic electric rates. *HVAC&R Research* 13(4):557-580.

Braun, J.E. 2007b. Impact of control on operating costs for cool storage systems with dynamic electric rates. *ASHRAE Transactions* 113.

Braun, J.E. 2007c. Near-optimal control strategies for hybrid cooling plants. *HVAC&R Research* 13(4):599-622.

Braun, J.E. 2007d. A general control algorithm for cooling towers in cooling plants with electric and/or gas-driven chillers. *HVAC&R Research* 13(4):581-598.

Braun, J.E. and G.T. Diderrich. 1990. Near-optimal control of cooling towers for chilled-water systems. *ASHRAE Transactions* 96(2):806-813.

Braun, J.E., J.W. Mitchell, S.A. Klein, and W.A. Beckman. 1987. Performance and control characteristics of a large central cooling system. *ASHRAE Transactions* 93(1):1830-1852.

Braun, J.E., S.A. Klein, J.W. Mitchell, and W.A. Beckman. 1989a. Applications of optimal control to chilled water systems without storage. *ASHRAE Transactions* 95(1):663-675.

DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*

Braun, J.E., S.A. Klein, J.W. Mitchell, and W.A. Beckman. 1989b. Methodologies for optimal control to chilled water systems without storage. *ASHRAE Transactions* 95(1):652-662.

Braun, J.E., K.W. Montgomery, and N. Chaturvedi. 2001. Evaluating the performance of building thermal mass control strategies. *International Journal of HVAC&R Research* (now *HVAC&R Research*) 7(4):403-428.

Brent, R.P. 1973. *Algorithms for minimization without derivatives*, Ch. 5. Prentice Hall.

Brothers, P.W. and M.L. Warren. 1986. Fan energy use in variable air volume systems. *ASHRAE Transactions* 92(2B):19-29.

Cheng, H., M.J. Brandemuehl, G.P. Henze, A.R. Florita, and C. Felsmann. 2008. Evaluation of the primary factors impacting the optimal control of passive thermal storage (RP-1313). *ASHRAE Transactions* 114(2):57-64.

Cumali, Z. 1988. Global optimization of HVAC system operations in real time. *ASHRAE Transactions* 94(1):1729-1744.

Cumali, Z. 1994. Application of real-time optimization to building systems. *ASHRAE Transactions* 100(1).

Drees, K.H. 1994. *Modeling and control of area-constrained ice storage systems*. M.S. thesis, Purdue University, West Lafayette, IN.

Drees, K.H. and J.E. Braun. 1995. Modeling of area-constrained ice storage tanks. *International Journal of HVAC&R Research* (now *HVAC&R Research*) 1(2):143-159.

Drees, K.H. and J.E. Braun. 1996. Development and evaluation of a rule-based control strategy for ice storage systems. *International Journal of HVAC&R Research* (now *HVAC&R Research*) 2(4):312-336.

Dyer, D.F. and G. Maples. 1981. *Boiler efficiency improvement*. Boiler Efficiency Institute. Auburn, AL.

Elleson, J.S. 1996. *Successful cool storage projects: From planning to operation*. ASHRAE.

Englander, S.L. and L.K. Norford. 1992. Saving fan energy in VAV systems, Part 2: Supply fan control for static pressure minimization using DDC zone feedback. *ASHRAE Transactions* 98(1):19-32.

Florita, A.R. and G.P. Henze. 2009. Comparison of short-term weather forecasting models for model predictive control. *HVAC&R Research* 15(5):835-853.

Forrester, J.R. and W.J. Wepfer. 1984. Formulation of a load prediction algorithm for a large commercial building. *ASHRAE Transactions* 90(2B):536-551.

Gayeski, N.T., J. Gagne, P.R. Armstrong, S. Katipamula and M. Alvira. 2011a. Development of a Low-Lift Chiller Controller and Simplified Precooling Control Algorithm–Final Report, PNNL-21155, Pacific Northwest National Laboratory, Richland, WA.

Gayeski, N.T., P.R. Armstrong, T. Zakula and L.K. Norford. 2011b. Empirical modeling of a rolling-piston compressor heat pump for predictive control in low-lift cooling, *ASHRAE Transactions 117 (2)*.

Gayeski, N.T., P.R. Armstrong, and L.K. Norford, 2012. Predictive pre-cooling of thermo-active building systems with low-lift chillers, *HVAC&R Research*, 18 (5), 858-873

Gibson, G.L. and T.T. Kraft. 1993. Electric demand prediction using artificial neural network technology. *ASHRAE Journal* 35(3):60-68.

Greensfelder, E.M., G.P. Henze, and C. Felsmann. 2011. An investigation of optimal control of passive building thermal storage with real time pricing. *Journal of Building Performance Simulation*, 4(2).Greensfelder, E.M., G.P. Henze, and C. Felsmann. 2011. An investigation of optimal control of passive building thermal storage with real time pricing. *Journal of Building Performance Simulation*, 4(2).

Grumman, D.L. and A.S. Butkus, Jr. 1988. Ice storage application to an Illinois hospital. *ASHRAE Transactions* 94(1):1879-1893.

Hackner, R.J., J.W. Mitchell, and W.A. Beckman. 1984. HVAC system dynamics and energy use in buildings—Part I. *ASHRAE Transactions* 90(2B):523-535.

Hackner, R.J., J.W. Mitchell, and W.A. Beckman. 1985. HVAC system dynamics and energy use in buildings—Part II. *ASHRAE Transactions* 91(1B):781.

Hartman, T. 1993. Terminal regulated air volume (TRAV) systems. *ASHRAE Transactions* 99(1):791-800.

Henze, G.P., M. Krarti, and M.J. Brandemuehl. 1997a. A simulation environment for the analysis of ice storage controls. *International Journal of HVAC&R Research* (now *HVAC&R Research*) 3(2):128-148.

Henze, G.P., R.H. Dodier, and M. Krarti. 1997b. Development of a predictive optimal controller for thermal energy storage systems. *International Journal of HVAC&R Research* (now *HVAC&R Research*) 3(3):233-264.

Henze, G.P., M. Krarti, and M.J. Brandemuehl. 2003. Guidelines for improved performance of ice storage systems. *Energy and Buildings* 35(2):111-127.

Henze, G.P., D. Kalz, S. Liu, and C. Felsmann. 2005. Experimental analysis of model-based predictive optimal control for active and passive building thermal storage inventory. *HVAC&R Research* 11(2):189-214.

Henze, G.P., M.J. Brandemuehl, C. Felsmann, A. Florita, and H. Cheng. 2007. Evaluation of building thermal mass savings. ASHRAE Research Project RP-1313, *Final Report*.

Henze, G.P., C. Felsmann, A.R. Florita, M.J. Brandemuehl, H. Cheng, and C.E. Waters. 2008. Optimization of building thermal mass control in the presence of energy and demand charges (RP-1313). *ASHRAE Transactions* 114(2):75-84.

Henze, G.P., A.R. Florita, M.J. Brandemuehl, C. Felsmann and H. Cheng. 2009. Advances in near-optimal control of passive building thermal storage. *Proceedings of the ASME 3rd International Conference on Energy Sustainability*, San Francisco.

Hiller, C.C. and L. Glicksman. 1977. Heat pump improvement using compressor flow modulation. *ASHRAE Transactions* 83(2)

Kao, J.Y. 1985. Control strategies and building energy consumption. *ASHRAE Transactions* 91(2B):510-817.

Katipamula, S., P.R. Armstrong, W. Wang, N. Fernandez, H. Cho, W. Goetzler, J. Burgos, R. Radhakrishnan, and C. Ahlfeldt. 2010a. *Cost-effective integration of efficient low-lift baseload cooling equipment: FY08 final report*. PNNL-19114.

Katipamula, S., P.R. Armstrong, W. Wang, N. Fernandez, H. Cho, W. Goetzler, J. Burgos, R. Radhakrishnan, and C. Ahlfeldt. 2010b. *Development of high-efficiency low-lift vapor compression system—Final report*. PNNL-19227.

Keeney, K.R. and J.E. Braun. 1996. A simplified method for determining optimal cooling control strategies for thermal storage in building mass. *International Journal of HVAC&R Research* (now *HVAC&R Research*) 2(1):1-20.

Keeney, K.R. and J.E. Braun. 1997. Application of building precooling to reduce peak cooling requirements. *ASHRAE Transactions* 103(1):463-469.

Krarti, M., M.J. Brandemuehl, and G.P. Henze. 1996. Evaluation of optimal control for ice systems (RP-809). ASHRAE Research Project, *Report*.

Krarti, M., G. Henze, G. Zhou, P. Ihm, S. Liu, and S. Morgan. 2007. *Real-time predictive optimal control of active and passive thermal energy storage systems: Final report*. Department of Energy Contract DE-FC-36-03G)13026.

Kreider, J.F. and X.A. Wang. 1991. Artificial neural networks demonstration for automated generation of energy use predictors for commercial buildings. *ASHRAE Transactions* 97(2):775-779.

Lau, A.S., W.A. Beckman, and J.W. Mitchell. 1985. Development of computer control—Routines for a large chilled water plant. *ASHRAE Transactions* 91(1B):766-780.

Liu, S. and G.P. Henze. 2006a. Experimental analysis of simulated reinforcement learning control for active and passive building thermal storage inventory—Part 1: Theoretical foundation. *Energy and Buildings* 38(2):142-147.

Liu, S. and G.P. Henze. 2006b. Experimental analysis of simulated reinforcement learning control for active and passive building thermal storage inventory—Part 2: Results and analysis. *Energy and Buildings* 38(2):148-161.

Ljung, L. and T. Söderström. 1983. *Theory and practice of recursive identification*. MIT Press, Cambridge, MA.

MacArthur, J.W., A. Mathur, and J. Zhao. 1989. On-line recursive estimation for load profile prediction. *ASHRAE Transactions* 95(1):621-628.

Morgan, S. and M. Krarti. 2006. Impact of electricity rate structures on energy cost savings of pre-cooling controls for office buildings. *Building and Environment* 42(8):2810-2818.

Morgan S. and M. Krarti. 2010. Field testing of optimal controls of passive and active thermal storage. *ASHRAE Transactions* 116(1).

Morris, F.B., J.E. Braun, and S. Treado. 1994. Experimental and simulated performance of optimal control of building thermal storage. *ASHRAE Transactions* 100(1):402-414.

Niswander, A., and P.R. Armstrong, 2013. Demand responsive cooling with TABS, *Intelligent Building Operation Conf.*, U. Colorado Boulder.

Niswander, A., 2014. *Demand Responsive Cooling with TABS Source/Sink and Model-Predictive Control*, MSME Thesis, Masdar Inst., UAE.

Rabl, A. and L.K. Norford. 1991. Peak load reduction by preconditioning buildings at night. *International Journal of Energy Research* 15:781-798.

Ruud, M.D., J.W. Mitchell, and S.A. Klein. 1990. Use of building thermal mass to offset cooling loads. *ASHRAE Transactions* 96(2).

Rawlings, L.K. 1985. Strategies to optimize ice storage. *ASHRAE Journal* 27(5):39-44.

Seem, J.E. and J.E. Braun. 1991. Adaptive methods for real-time forecasting of building electrical demand. *ASHRAE Transactions* 97(1):710-721.

Seem, J.E., S.A. Klein, W.A. Beckman, and J.W. Mitchell. 1989a. Comprehensive room transfer functions for efficient calculation of the transient heat transfer process in buildings. *Journal of Heat Transfer* 111:264-273.

Seem, J.E., P.R. Armstrong, and C.E. Hancock. 1989b. Comparison of seven methods for forecasting the time to return from night setback. *ASHRAE Transactions* 95(2).

Seem, J.E., S.A. Klein, W.A. Beckman, and J.W. Mitchell. 1990. Model reduction of transfer functions using a dominant root method. *ASME Journal of Heat Transfer* 112:547-554.

Seem, J.S., C. Park, and J.M. House. 1999. A new sequencing control strategy for air-handling units. *International Journal of HVAC&R Research* (now *HVAC&R Research*) 5(1):35-58.

Snyder, M.E. and T.A. Newell. 1990. Cooling cost minimization using building mass for thermal storage. *ASHRAE Transactions* 96(2):830-838.

Spethmann, D.H. 1989. Optimal control for cool storage. *ASHRAE Transactions* 95(1):1189-1193.

Stoecker, W.F. 1980. *Design of thermal systems*. McGraw-Hill, New York.

Tamblyn, R.T. 1985. Control concepts for thermal storage. *ASHRAE Transactions* 91(1B):5-11.

Taplin, H.R. 1998. *Boiler plant and distribution system optimization manual*. Fairmont Press, Lilburn, GA.

Warren, M. and L.K. Norford. 1993. Integrating VAV zone requirements with supply fan operation. *ASHRAE Journal* 35(4):43-46.

Watkins, C. and P. Dayan. 1992. Q-learning. *Machine Learning* 8:279-292.

Wei, G., M. Liu, and D.E. Claridge. 2004. Integrated damper and pressure reset for VAV supply air fan control. *ASHRAE Transactions* 110(2):309-313.

Winkelman, F.C., W.F. Buhl, B. Birdsall, A.E. Erdem, K.L. Ellington, and Hirsch & Associates. 1993. DOE-2 BDL summary, version 21.E. LBNL *Report* 34946, Lawrence Berkeley National Laboratory, Berkeley, CA.

Zakula, T., N.T. Gayeski, P.R. Armstrong, and L.K. Norford. 2011. Variable-speed heat pump performance model, model validation, and optimal control. In preparation for submission to *HVAC&R Research.*17(5):670-691

Zakula, T., P Armstrong, L Norford, 2012. Optimal coordination of heat pump compressor and fan speeds and subcooling over a wide range of loads and conditions, *HVAC&R Research* 18 (6), 1153-1167

Zakula, T., 2013. *Model-Predictive Control for Energy Efficient Cooling and Dehumidification*, PhD Thesis, MIT.

Zakula, T., P.R. Armstrong, L.K. Norford. 2014a. Modeling Environment for Model Predictive Control of Buildings, ENB-D-14-00483, *Energy and Buildings*

Zakula, T., P.R. Armstrong, L.K. Norford. 2014b. Advanced Cooling Technology with Thermally Activated Building Surfaces and Model Predictive Control, ENB-D-14-00484, *Energy and Buildings*

Zhou, G., P. Ihm, M. Krarti, S. Liu, and G.P. Henze. 2005. Integration of an internal optimization module within EnergyPlus. *IBPSA Proceedings*.

DRAFT: 2015 ASHRAE Handbook—*HVAC Applications*