Control Challenges and Opportunities in Building Automation

Scott A. Bortoff, Bryan Eisenhower, Veronica Adetola, Zheng O'Neill

30-October-2023

Abstract

Building automation systems provide supervisory control of a building's heating, ventilation and air conditioning (HVAC), electrical, lighting, shading, access control elevators and escalators, security systems, etc., primarily for purposes of automation (reducing manual labor), reducing energy consumption, improving occupant comfort, and recording and communicating information for maintenance and accountability. Control as a subject can drive meaningful innovation in building automation, particularly to reduce the impact that buildings have on climate such as greenhouse gas production. HVAC systems in particular, as the largest consumer of power in buildings, are strongly affected by control, and continued efforts by the industry to improve energy efficiency have made them increasingly multivariable, dynamically interactive, and nonlinear. In this chapter, we focus on both need-based and vision-based control innovation for building automation for HVAC systems. After providing a brief background that describes different HVAC architectures and products, we describe the control challenges and opportunities from a needs-based point of view, at the equipment level, the systems level, and the building level. At the equipment level, increased use of continuously variable actuation such as variable speed compressors and fans, requires application of robust multivariable control that must consider various types of nonlinear behavior across increasingly larger operating envelopes. The objective at this level is to achieve robust and energy efficient operation. At the system level, coupled and interactive dynamics among subsystems are increasingly important to consider, and functional integration of these subsystems, along with set points and operating schedules are designed to minimize energy consumption and improve occupant comfort. At the building level, issues such as integration among other building systems, such as the building envelope and the electric grid, offer opportunities to develop new types of demand response and grid-interactive behaviors and technologies. At the higher levels, system dynamics and robust stability are less of an issue, similar to process control applications, while optimization and integration are more important, although dynamic stability is increasing an issue at the higher levels too. From a vision-driven innovation point of view, we describe three areas that will require sustained research efforts and will eventually impact building automation: Digital Twins, Model Predictive Control, and Grid Interactive Buildings. Control as a field can and must play an active role in developing next generation technologies at all levels of building automation, especially because these systems are increasingly dynamically interactive and robust stability is increasingly an important design consideration. These issues, and especially the recognition of the relationship between robustness and performance, are central to control as a field, and arguably differentiate it from related research fields that may emphasize only performance.

Building automation is commonly defined as the automatic centralized control of a building's heating, ventilation and air conditioning (HVAC), electrical, lighting, shading, access control, elevators and escalators, security systems, and other interrelated systems through a Building Management System (BMS).[1] The primary purpose of a BMS is to automate certain aspects of these systems' operations, reducing or eliminating the need for manual labor. The intentions are to reduce energy consumption, to increase security, to improve occupant comfort, to record information for maintenance or accountability reasons, and to provide remote access to this information.

In this chapter we pose the question: What are the control research challenges that, if addressed, will enable and drive meaningful innovation in building automation in the 21st century? We focus on innovations that might mitigate the impact that buildings have on the climate, especially greenhouse gas production. Buildings, as end users of various forms of energy (mainly electricity, natural gas and oil), account for approximately one third of greenhouse gas emissions worldwide. Building automation systems therefore have a large potential to directly reduce emissions, through improved conservation and efficiency measures, and also to indirectly reduce or eliminate emissions, by enabling behaviors such as demand response (DR), for example, that in turn enable innovations in the supply of renewable electricity.

In this chapter we limit our scope to building automation as it pertains to HVAC systems and associated building systems such as active shading, photovoltaics (PV) and electric grid-interaction, for two reasons. First, HVAC systems are the largest consumers of energy in commercial and residential buildings [6, 1]. Second, performance and behavior of HVAC systems are strongly affected by control, whereas other systems associated with a BMS, such as security or access control, are information systems that are generally outside the influence of control, and/or have relatively little impact on energy consumption. We also limit discussion to common types of residential and commercial buildings, including single or multi-family residential buildings, apartments, office buildings, schools, hospitals, hotels, restaurants, retail buildings and the like. Some types of buildings have specialized energy or HVAC requirements, such as those used for industrial manufacturing, and may fall outside our scope. Other types of buildings such as supermarkets have significant amounts of refrigeration equipment, which makes use of the same vapor compression principles as HVAC. However, for simplicity of exposition, we ignore the specialized needs and research challenges associated with these buildings, although some of the issues we raise, and some of the needs for innovation that we describe, may be applicable to these types of buildings as well.

On the other hand, we adopt a broader definition of building automation than the common one provided above, to include all aspects of the operation, automation and control of HVAC systems and equipment (and related systems) within buildings, not only a centralized BMS, which typically provides only a supervisory level of control. This is also for two reasons. First, the overall system performance is affected by all levels of control, and non-trivial challenges and opportunities exist even at the lowest level. Second, HVAC systems manifest a very broad variety of physical manifestations, depending on building type and location. Some are so-called "built up" systems, which are designed and constructed as a custom system for a specific, individual building. Other buildings employ factory-built, packaged unitary or split equipment, where the control function is inserted into the product at the factory, designed to be commissioned and integrated into a broad variety of building types. The control of this type of equipment increasingly includes functionality that is conventionally associated with a BMS, blurring the hierarchal organization. So we consider all levels of building HVAC control in buildings.

This chapter is organized as follows. In Section 1 we provide a brief description of a typical vapor compression cycle, which is by far the most common and efficient means to move heat in buildings.

[1] https://en.wikipedia.org/wiki/Building_automation

We then describe several common types of HVAC systems used in commercial and residential buildings, which are typically organized into a hierarchal structure ranging from equipment at the lowest level to the building or a collection of buildings at the highest. We then describe dynamic models that are used for model-based control development. In Section 2 we discuss industry trends and drivers of innovation, followed by Section 3 which articulates their implications on control. In Section 4 we describe opportunities for innovation and control research from the need-driven point of view of the industry. In Section 5, we imagine a set of vision-driven opportunities for innovation and control research. We conclude the chapter in Section 6, returning to the question: How can the control research community impact building automation?

1 HVAC Background

Building HVAC systems provide two functions to buildings: 1) Regulation of thermal comfort, typically by regulating indoor air temperature and possibly humidity, and/or regulating certain building construction temperatures, and 2) ventilation, by actively exchanging some of the inside air with outside air. (Except for extreme climates, residential HVAC systems typically provides only (1), with (2) being passively achieved by air infiltration through the building envelope.) In order to meet (1), heat needs to be moved, usually from a colder fluid to warmer fluid, requiring energy to be consumed. In order to meet (2), air must be moved from one location to another. The basic technologies for both has remained the same for more than a century: the vapor compression cycle and fans.

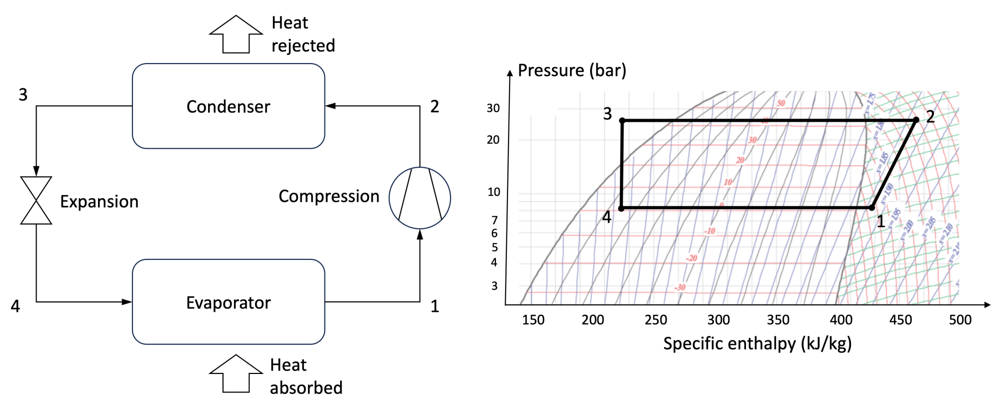

Figure 1: Basic vapor compression cycle (left) with corresponding pressure-enthalpy chart (right).

1.1 Vapor Compression Cycle

At the heart of nearly every building HVAC system is a closed vapor compression cycle, diagrammed in Figure 1, whose purpose is to move heat from a lower temperature fluid to a higher temperature fluid using mechanical (or in some cases chemical) input power. Other physical processes exist for this purpose, such as the thermoelectric effect or magnetic refrigeration, but neither comes close to the energy efficiency and scalability of vapor compression. The basic vapor compression cycle operates as follows. A compressor compresses the refrigerant to hot, super-heated gas, which flows into the condenser heat exchanger (HEX) where it condenses, releasing heat to the first fluid stream. Upon leaving the condenser HEX, the refrigerant is expanded through an expansion device, where its pressure (and temperature) drops, but, to first approximation, no heat transfer occurs. At the exit of the expansion device, the refrigerant is typically two-phase (part gas and part liquid), and it enters the evaporator HEX, where it evaporates, absorbing heat from the second fluid stream. Usually it evaporates entirely to a slightly superheated gas at the exit of the evaporator HEX, where it is sucked into the compressor, completing the cycle. Thermodynamically the cycle is diagrammed using the pressure-enthalpy chart on the right of Figure 1, with numbers corresponding to points between the major components.

A practical vapor compression system usually includes additional components, such as accumulators or receivers, which store liquid refrigerant at various locations in the cycle, valves to change the direction of the refrigerant flow, so that the system can move heat in either direction, and additional HEXs to improve thermodynamic efficiency, and may also incorporate multiple compressors, or HEXs on either side of the cycle, as shown in Figure 2.

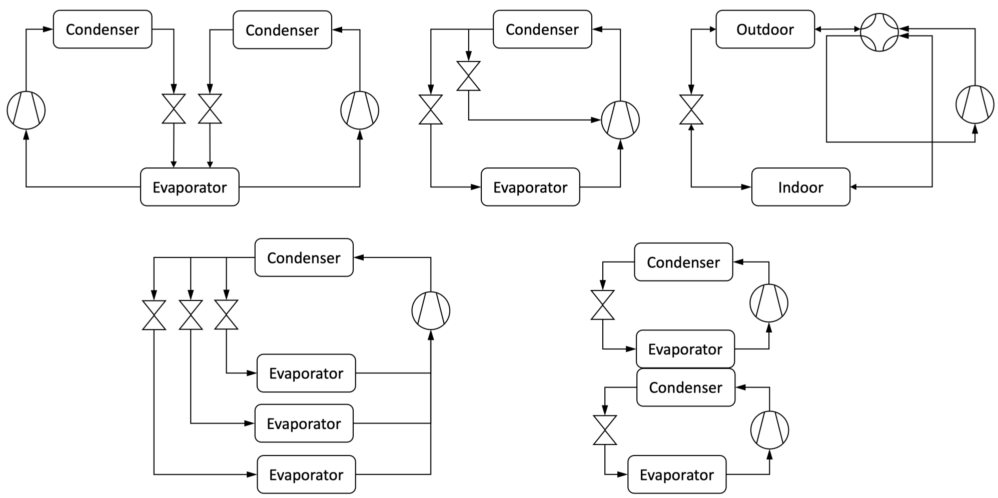

Figure 2: Non-exhaustive examples of different vapor compression system architectures (clockwise); dual circuit, economized, reversible (heat pump), multi-evaporator, and two-stage.

1.2 Fans

Fans remain the basic means of air transport in buildings, and come in a variety of types, including centrifugal, axial and cross-flow (roll). Although little has changed in the basic principles of operation during the last century, modern fan design and control is increasingly complex, with energy efficiency, operating envelope (due to variable-speed actuation) and acoustic noise all being important design factors, and coupling and dynamic interaction among components of an overall system affecting its controlled operation.

1.3 HVAC System Architectures

Building HVAC systems may be classified depending on the building type, size, and location. At the largest end of the spectrum, a large commercial building (or campus) typically will have a so-called “built-up” Variable Air Volume (VAV) HVAC system, consisting of a chilled water plant (chillers, water distribution, cooling towers), air handler units (supply and return air fans, heating and cooling HEXs, filters), air distribution ducts, and terminal units with controlled dampers for

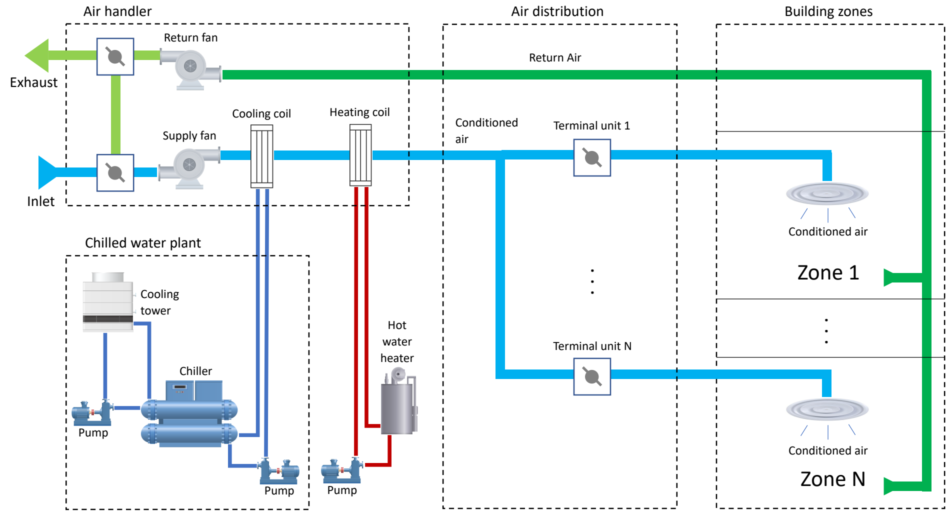

Figure 3: Built-Up System.

each building zone, as shown in Figure 3. The chiller uses a vapor compression cycle to move heat from the chilled water loop to the higher temperature condenser water loop. What characterizes this type of system is its use of water as a means to transport heat over large distances, and the ducted distribution of conditioned air throughout the building. This type of system is custom designed for the building and assembled on-site from factory-made components.

Figure 4: A packaged unitary HVAC product.

Smaller commercial buildings are typically served by so-called packaged unitary HVAC equipment, in which the function of conditioning the air is packaged into a factory-built unit, often placed on the roof of the building. An example is shown in Figure 4. Packaged unitary equipment will typically include supply and return fans for ventilation, heat exchangers, a vapor compression system, and possibly gas or electric heat, depending on the application. Usually the vapor compression system is air-source, meaning heat is exchanged directly with the outside air on one side of the vapor compression cycle, while the other conditions the indoor air stream, so refrigerant is used to move heat directly from the indoor air to the outdoor air. (These systems are sometimes called "direct expansion," or DX units.) As in Figure 3, the conditioned air is distributed throughout the building, and local zone conditions are regulated by feedback by actuating a variable-position damper in a terminal unit.

A third architecture, commonly used in Europe and Asia, is based on so-called Variable Flow Refrigerant (VRF) equipment. As shown in Figure 5, the vapor compression cycle is physically split

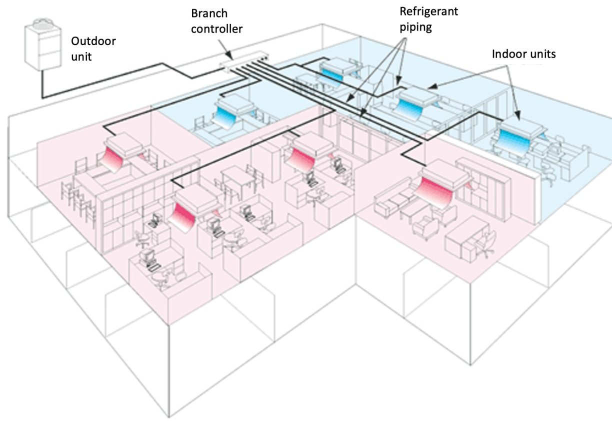

Figure 5: A VRF system.

into two or more separate units (so these are sometimes called "split systems"). The outdoor unit contains the compressor, a heat exchanger, variable speed fans and valves, and possibly refrigerant expansion devices. Indoor units are located throughout the building and contain a heat exchanger, a fan and possibly refrigerant expansion devices. Pipes connect the indoor and outdoor units, and in some systems, additional equipment such as branch controllers are used between the two. These systems usually have variable speed compressors, so that the cooling / heating capacity (and the refrigerant flow rate) can be varied. Ventilation is supplied via an Outside Air Processor (OAP) or Dedicated Outdoor Air System (DOAS), which is a separate unit that is similar to a packaged unitary HVAC system, but smaller because it conditions only the outside air that is introduced into the building. These may incorporate energy exchange between the supply and exhaust air, such as desiccant wheels, for energy saving. What characterizes these systems is that they transport heat through the building using refrigerant (instead of water), and they regulate zone temperatures by varying heat flux from each indoor unit. Therefore VRF systems do not require most of the air ducting used in VAV systems, and are sometimes called "duct-free." VRF systems have an advantage that in some configurations, they can move heat from one building zone to another (for example, moving heat from one room in cooling mode to another in heating mode), which is thermodynamically more efficient than exchanging both with the outside air.

For single-family residential building application, many systems in North America are so-called forced-air architectures with a split vapor compression system. The outdoor unit includes the compressor while the indoor HEX is mounted in a supply air duct. Depending on climate, this may be integrated to a gas-burning furnace, which provides heating in winter. In some systems, the vapor compression system can be driven as a heat pump, reducing the need for a furnace. Hydronic systems, which heat water using a vapor compression cycle and distribute it among indoor HEXs, are used for heating throughout much of Europe and Asia. Similarly, split VRF systems are also used for residential application, mainly in Europe and Asia, and increasingly in North America.

A myriad variations of these three basic architectures exist and this short description is non-exhaustive and incomplete [2, 3]. In some commercial buildings, terminal units incorporate heat exchangers that reheat conditioned air. Fan coil units, which include an air-to-water HEX and fan, and modulate heat flux to a building zone using either a water valve or the fan speed. In these systems, chilled and heated water is circulated throughout the building, instead of only the air handlers. In many parts of the world, district heating systems circulate warm water or steam through a piping network to buildings, which then contain their own hydronic heating system. Some buildings use radiant heating and cooling, in which heated or chilled water is circulated through parts of building constructions (usually the floor or ceiling), and heat is transferred convectively and radiatively from these surfaces to building occupants. Hybrid VRF systems are a mix of split VRF systems and hybrid systems; the vapor compression cycle is split, and indoor HEXs exchange heat with a water loop, which distributes heat among a set of terminal units, usually fan coils. Many buildings include a mix of these different architectures, because over the building lifecycle, heating and cooling requirements evolve. For example, computer servers rooms require cooling year round, with high heat flux requirements. Often this is met using a dedicated VRF system even though the building may be an entirely different architecture.

1.4 Control

Regardless of the architecture, control of these system is hierarchal, resembling a process control application, although the physical manifestations of the control functions vary by HVAC architecture and equipment type. It can be described as having three layers, although this is a somewhat arbitrary partition. At the lowest levels, feedback control loops regulate process variables to set points and enforce process variable constraints. Many control loops are SISO, with PID compensation if the actuator is continuously variable, or with hysteresis if the actuator is on-off. At the system level, the control issues concern coordination and functional integration, so that components function together as an automated system or subsystem. At this level, start-up and shut-down sequences are realized, some set-points are scheduled, and some protection logic, error handling, and alarm logic is realized. At the highest level, where the conventional BMS sits, building-level issues are automated, and integration and coordination among building-level systems such as fire and security are realized. Building occupancy schedules are defined, resets (set-point changes to minimum-energy settings) are realized, and data is collected, marshaled and communicated for purposes of diagnostics and remote monitoring.

1.4.1 Low Level Control

At this level, the primary requirements are regulation, disturbance rejection and constraint enforcement of process variables. What characterizes this level of the hierarchy is the use of feedback. Although many feedback loops are SISO, the entire HVAC system and building are dynamically coupled and highly interactive, and robust stability and performance over wide operating ranges and in a variety of building applications, as the equipment and building evolve, is the primary challenge. The quintessential example is zone temperature regulation. In one of the simplest and oldest examples of feedback, a wall-mounted thermostat turns on or off the heating or cooling equipment depending on the difference between zone temperature set point and measured zone temperature. This is still the most common method of comfort control in residential buildings today.

VAV systems have many interacting feedback loops. Each building zone temperature is regulated by actuating a continuously variable damper located inside the zone's terminal unit, often via a PID compensator. Upstream, the duct air pressure is regulated by feedback to the supply air fan. When the temperature increases in a zone, the dampers open to introduce more conditioned air into the zone. This reduces the duct pressure, and the pressure loop responds by increasing the supply air fan speed to maintain the pressure. The ratio of outside air to recycled air may also be regulated by feedback using a CO_2 sensor. If present, the cooling tower return water temperature is regulated by feedback, typically by varying the cooling tower fan speed. These functions are realized in an embedded digital controller, which often controls multiple zones and interfaces to higher levels of control, and is configured (tuned) at commissioning time, as the building is being constructed or modified.

Many low-level control functions lie within HVAC equipment, such as chillers or VRF systems, and are designed at the factory. These so-called Product Inserted Controls (PICs) contain proprietary control algorithms that have, over time, become increasingly complex and sophisticated. For example, in some types of chillers, the basic control actuators are the compressor speed, which can be changed continuously, and an electronic expansion valve setting, which is also continuously variable. (Additionally, chillers that use a centrifugal compressor also control aspects of refrigerant flow into the compressor, usually using a continuously-variable inlet guide vane.) If the water pumps are integrated, then their speeds may also be under control, either as continuously variable or as on-off. The basic process variables to be regulated in the chiller are the leaving water temperature and the evaporator superheat. Constraints on water temperatures, flow rates, refrigerant pressures and various variables related to the compressor, depending on its physical type, need to be enforced, typically by selector and anti-windup logic. The chiller PIC will integrate start up and shut down sequences, which ramp up down and down actuation and close feedback loops, and also various forms of protection logic that prevents equipment damage. It may also include logic to generate alarms, some diagnostics, and functionality to record data for maintenance uses, e.g. runtime, number of start-stop cycles, etc.

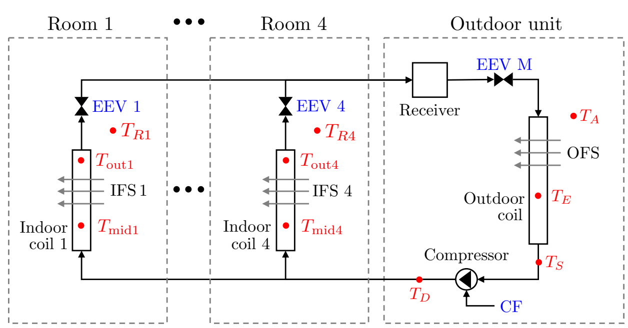

Figure 6: Four-zone VRF system showing temperature sensors (red) and control actuators (blue) [12]. Refrigerant flows clockwise in heating mode and counterclockwise in cooling mode, changed by a valve not shown.

The PIC for a Variable Refrigerant Flow (VRF) system regulates and enforces constraints on internal process variables, and additionally provides for a significant amount of building-level functionality, because as a system, the factory-built VRF provides many of the functions of the building HVAC. For example, consider the four-zone House Air Conditioner (HAC) system diagrammed in Figure 6. Even though this is a relatively small VRF system, it has ten continuously variable control actuators: The compressor frequency, the commanded settings for five electronic expansion valves (EEVs), four indoor fan speeds, and the outdoor unit fan speed. It can run in either heating or cooling mode. It regulates room temperatures at the building level, allowing for different zone temperatures, and also regulates a set of internal process variables while also enforcing several process constraints using a multivariable PID with anti-windup and selector and override logic [13].

1.4.2 System Level Control

At the system level, the primary requirements are coordination, functional integration among components and subsystems, and automation. At this level, the operation of separate components is coordinated in order to achieve a set of system level performance requirements. For example, the chillers, water pumps, valves and cooling towers all need to work together in order to deliver chilled water to the air handlers. Part of this is to automate start-up and shut-down sequences, which can be realized in a programmable logic controller (PLC) using finite state machines or ladder logic. For example, a large chilled water plant usually includes several different types of chillers, with one being energy efficient across a wide operating range but having relatively small capacity, while the others may be larger and most efficient at their highest capacity. (Typically the former employ screw-type compressors, while the latter use centrifugal type.) As building heat demand varies, the chillers need to be sequenced in the most energy efficient manner, so that the larger chillers run at maximum capacity. A supervisory control will encode the logic to sequence the controllers, and will also rotate their individual operation to take into account maintenance issues.

Some set points for the lower-level feedback loops are determined at the system level, often by a schedule that might depend on measured building conditions or time-of-day and occupancy schedules. For example, the set point for the leaving chilled water might be scheduled on outside air conditions, with colder set points corresponding to warmer, humid weather. The colder chilled water will make the HEX colder, which condenses more water, lowering indoor humidity on hot, humid days. So-called resets are included at this level, which modify process variable set points such as flow rates or supply temperatures to minimum values in order to save energy when the building is in an unoccupied state.

Although feedback loops in the conventional sense are generally considered at the lower level, stability, robustness and system dynamics are also key issues at this level because HVAC components and subsystems, when interconnected, become dynamically coupled and interactive. Emergent, system level dynamic behaviors can manifest themselves, and, in the limit, do so as undesirable oscillations or outright instability. Furthermore, the building's thermodynamics are tightly coupled to the HVAC equipment's dynamics, and this makes each installation different. Therefore, although most of the design considerations at this level concern automation, significant attention must be paid to overall system dynamics, stability, and robustness.

1.4.3 Building Level Control

At the building level of control, control is largely supervisory in nature. The primary issues are set-point scheduling, data collection and remote monitoring, and integration among other information systems such as lighting control, fire and security. Integration of physical systems such as the power supply (electric grid), local electrical generation, energy storage and the building envelope, are also considered. At this level, system dynamics are less of an issue, and optimization of performance and information integration are the primary considerations.

Zone temperature set-points are determined, usually using a time-of-day schedule, possibly combination with occupancy sensing. Some HVAC set points such as the supply air or water temperatures may also be determined at this level. A primary objective is performance monitoring, making visible key metrics of building operation, and data analytics are playing an increasingly important role. Some types of diagnostics and fault detection are deployed at this level. Electric power grid integration is also consideration. Building-level energy consumption can be shifted from high demand periods to lower demand periods by modifying zone set points. Demand Response (DR) allows the grid operator to reset zone temperature set points (for example, 5°F for 5 hours), reducing HVAC power consumption during peak demand periods via communication standards such as Organization for the Advancement of Structured Information Standards (OASIS) Energy Interoperation [28]. Building envelope integration is also considered at this level. Some buildings have active shading or automatically operable windows, which may be actuated by the BMS depending on time of day, weather etc.

Networking plays a large role at the building-level. Open standards such as BACnet are widely supported and used to collect and disseminate data on standard networking platforms. (However, at the systems level, networking is commonly done using standards that are proprietary to particular manufacturers, with translation provided to open standards at the building level.) BACnet allows for interoperability among a wide range of building automation applications, including HVAC, lighting, fire detection and access control.

1.5 Dynamic Models

Dynamic models of HVAC systems are used for several purposes, ranging from equipment-level control design to building-level energy optimization. The physics that govern the behavior of HVAC systems include heat transfer (conductive, advective and radiative) and fluid transport. Very generally speaking, the dynamics can be represented mathematically as a set of hybrid (including both continuous and discrete states) differential and algebraic equations (DAEs),

$$0 = f_x(v(t,t_-),\theta) \tag{1a}$$
$$\xi(t) = f_\xi(v(t,t_-),c(t),\theta) \tag{1b}$$
$$c(t) = f_c(v(t,t_-)) \tag{1c}$$
$$y(t) = h(x(t),z(t)). \tag{1d}$$

Here $v(t,t_-) := [\dot{x}(t), x(t), z(t), \xi(t), \xi(t_-)]$ is a vector of continuous and discrete states, $x(t)$ is a vector of differential states (refrigerant pressures and enthalpies, metal temperatures at discrete spatial locations in the vapor compression cycle, water temperatures and flow rates, building construction temperatures, zone air temperatures, pressures and humidities etc.), $z(t)$ is a vector of algebraic variables, $y(t)$ is a vector of measurements, $d(t)$ represents time-varying boundary conditions or disturbances that are typically not measured, $u(t)$ is a vector of continuous and discrete control actuators, $\xi(t)$ is a vector of discrete states (real, integer and boolean), which change only at discrete time events t_e and are otherwise constant, $c(t)$ is a vector of boolean states computed by evaluating relations e.g., $x_1 < x_2$, which change only at discrete time events t_e and are otherwise constant, and θ is a vector of physical parameters. The subscript t_- denotes time immediately prior to an event at time t, which is needed to define discrete event dynamics. Although the specific structure of (1) depends on the HVAC equipment type and architecture, building type and location, use case and modeling assumptions, there are certain general characteristics that are more-or-less universal and are important to consider from a control design and analysis point of view.

1.5.1 Large Scale, Nonlinear, Sparse

Depending on the application, (1) tends to be nonlinear and large-scale, with hundreds to perhaps millions of states. This is because heat exchangers are often modeled using finite-volume methods, in which each HEX is divided into a number of segments, with a refrigerant stream, metal walls, and an air (or water) stream associated with each segment. The refrigerant stream can be considered a one-dimensional fluid flow (gas, liquid or two-phase) and energy, mass and momentum balance equations are expressed for each segment. These are coupled to energy balance equations for the air (or water) stream and heat exchange at the metal wall boundaries, and augmented with a set of empirical closure relations describing single and two-phase heat transfer coefficients and frictional pressure drops. Taken together, each HEX can generate several hundred highly coupled, nonlinear differential algebraic equations, depending on its size, circuiting and type. Compressors are often modeled with a set of coupled, nonlinear algebraic equations that relate the compressor speed,

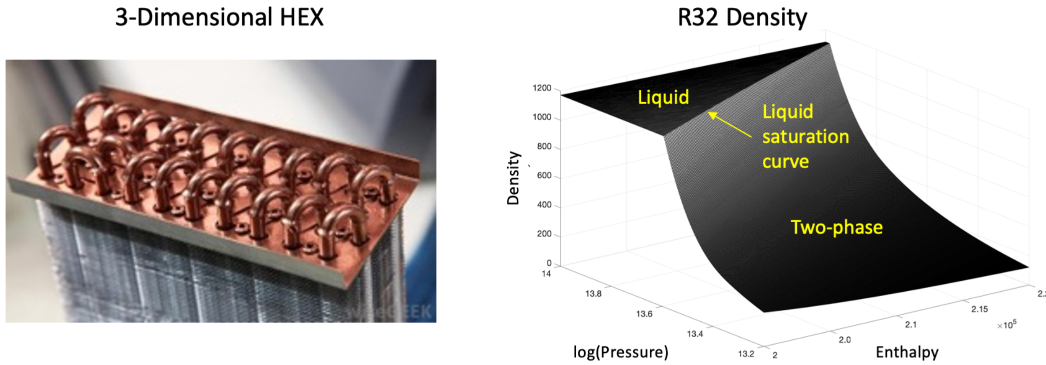

Figure 7: Multi-row air-source tube-fin heat exchanger (left), and the density of R32 as a function of enthalpy and pressure (right), showing strong nonlinearity along the liquid saturation curve, which separates liquid (left) from two-phase (right).

power consumption and thermofluid boundary states to one another. Expansion valves are also modeled algebraically, as are other components such as water pumps and fans.

These component models are composed into a system model to represent the interconnections among components, and coupled to a set of thermofluid dynamic equations that describe air transport, psychometrics and heat transfer with the building airspace, envelope and outdoor environment. Building envelope constructions are often modeled as layered, with conductive heat transfer between layers, radiative heat transfer among constructions, and convective heat transfer to the indoor air. The building model states include the temperature of each solid material layer, along with pressure, temperature and humidity of air in each zone. A fully composed dynamic system model typically has hundreds to tens of thousands of states (the dimension of x), and thousands to perhaps millions of algebraic variables (the dimension of z), depending on the size of the building, HVAC architecture and model resolution assumptions.

Nonlinearity arises in part because of refrigerant and water phase change. The balance equations for the refrigerant stream include partial derivatives of refrigerant density with respect to other fluid variables, which varies by several orders of magnitude depending on the refrigerant state (see Figure 7). Heat transfer is a strong nonlinear function of refrigerant state, with significantly higher heat transfer associated with two-phase refrigerant, compared with single-phase. Pressure - flow relationships are also nonlinear. Other nonlinear effects occur because of sensor placement, and will be described in Section 4.

Fortunately, composed HVAC system models tend to be sparse: Each equation in (1a) depends on a small number of states as a consequence of the finite volume modeling assumptions and the principle of locality. This holds from the small length scales of a HEX up to the large length and volume scales of a building. Figure 8 shows the sparsity pattern of the four-zone VRF system diagrammed in Figure 6: The Jacobian of (1) is less than 1% non-zero. This implies that specialized methods of numerical analysis for causalization, tearing and index reduction [17] are effective to generate control-oriented models and also efficient simulation code.

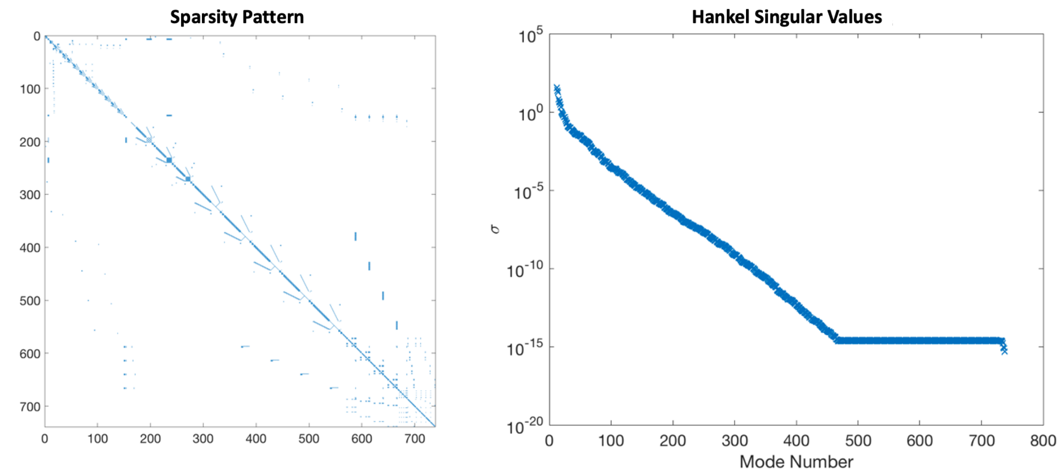

Figure 8: Sparsity pattern of the Jacobian for the four-zone VRF system diagrammed in Figure 6.

1.5.2 Numerically Stiff

On the other hand, the dynamic system (1a) is very numerically stiff. Time constants can range ten orders of magnitude. The fast dynamics are associated with fluid flow in the relatively small heat exchanger volumes, whereas the slow time constants are associated with larger, massive building constructions. Figure 9 plots the eigenvalues of the four-zone VRF system diagrammed in Figure 6, with the real axis on a logarithmic scale. The spectrum ranges from about 10μs to 1 day, or almost ten orders of magnitude. Numerical stiffness requires the use of implicit numerical solvers such as DASSL [14] when computing numerical solutions to (1), and can present challenges for model reduction, control and state estimator design.

It is important to emphasize that the dynamics couple when the HVAC system components are composed into a system model, and this is coupled to a building model. It is incorrect to view the "fluid pressure dynamics" in the vapor compression system as fast. In fact, refrigerant pressure as a state includes both fast and slow modes, and both are excited in normal operation of the system.

1.5.3 Hybrid

Finally, (1) is a *hybrid* model, containing continuous states x and variables z, and also discrete states ξ. HVAC systems include discrete actuators such as on/off valves which allow for multiple modes of operation. Control systems include discrete event subsystems to realize start-up, shut-down sequences, for example, and other types of logic. Fluid flow models also require discrete states that represent the flow direction; the structure of the continuous part of the equations depends on these states. For some aspects of control design, the operational mode can be considered constant, so that (1b)-(1c) may be neglected and the plant is treated as purely continuous-time. For the discrete-event aspects of control design, such as the discrete event start-up or shut-down sequences, which are usually represented as finite state machines, (1a) might be simplified, and parts of (1b)-(1c) are synthesized. At some point in the design process, the full hybrid model behavior must be considered, especially when switching among modes. For example, HVAC systems often operate in an on-off cycle when heat loads are low, and also switch from heating mode to defrost mode and back. Thus for all aspects and stages of control system design and simulation, multiple models are used to capture the relevant behavior.

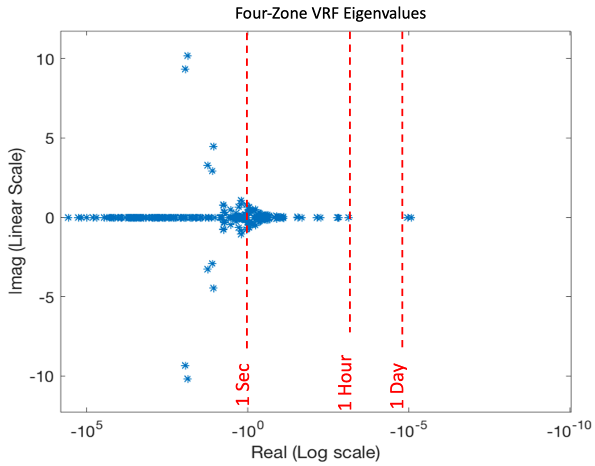

Figure 9: Eigenvalues of the Jacobian for the four-zone VRF system diagrammed in Figure 6, with real-axis on a logarithmic scale.

1.6 CAD Tools

Although (1) is a concise mathematical representation of an HVAC dynamic model, it offers little practical usefulness. In practice, Computer Aided Design (CAD) tools such as EnergyPlus, TRNSYS, Matlab or Modelica are used to construct, manage, analyze (to some limited extent) and simulate HVAC system dynamic models. This is necessary because of their size, complexity and numerical properties.

Particularly noteworthy is open-source *Modelica*[2][23, 34], which is a computer language for complex, multi-physical, heterogeneous systems such as HVAC in buildings. It is equation-based so that mathematical equations — the language of physics — may be transcribed naturally into the language. It is also object-oriented for organization, so that large models of complex, hierarchal systems, such as building HVAC, may be assembled from libraries of components, such as the open-source Buildings Library [46]. The hierarchal structure of the model mimics the structure of the physical system, and the use of object orientation facilitates reuse, so that component and system libraries accrue through use and grow into valuable business assets. Importantly, libraries of electric power generation and distribution e.g. PV, are available to meet growing needs to represent these interactions. The language is supported by commercial and open-source tools such as Dymola and OpenModelica [24], which compile a Modelica model into efficient executable simulation code, and can also compute a numerical or symbolic linearization about an equilibrium solution, which is vital for model-based control design.

What separates Modelica from alternatives is its hybrid differential-algebraic model of computation (in contrast, Matlab Simulink has a causal signal-flow model of computation), and the analysis steps that compilers automate to transform an acausal Modelica model into efficient, causalized simulation code. DAEs are more appropriate than ODEs for representing physics, especially compared to causal signal flow models, because they relieve the modeler from the need to make a priori assumptions about causality (which variables affect which other variables), which is error-prone and limits model applicability. Instead, the compiler determines this based on the system structure. This is especially important for thermo-fluid systems, because the flow direction changes the

[2]https://modelica.org/

structure of the equations, and cannot always be determined a priori. The support of hybrid DAEs (meaning a combination of both discrete event logic and continuous variables) enables rigorous implementation of synchronous objects such as flow networks in which fluid flow direction cannot be assumed a priori, and also digital control algorithms. The language includes synchronous language features to precisely define and synchronize sampled-data systems, including periodic, non-periodic and event-based clocks [21]. A real-time library allows for model interface to real-time input and output [42], allowing for experimental testing of control systems without the need for recoding.

2 Industry Trends and Drivers of Innovation

Several powerful forces are driving the global HVAC industry toward products that are increasingly complex, dynamic, interactive and functionally integrated. Here we summarize those most relevant.

2.1 Building-Level Energy Efficiency

Since 1975, ASHRAE Standard 90.1, *Energy Standard for Buildings Except Low-Rise Residential Buildings*, has evolved, defining a minimum acceptable standard for energy efficient design of buildings (except low-rise) in the US. The standard covers many building elements including the building envelope, HVAC, hot water, lighting, refrigeration equipment and on-site electrical generation. This standard bans certain types of energy-inefficient HVAC systems, such as constant air volume systems, while providing guidelines and minimum operating standards for other types of HVAC systems, such as variable air volume (VAV) systems. Most US states apply the standard or equivalent standards for all commercial buildings. Similar building standards exist in Asia and Europe.

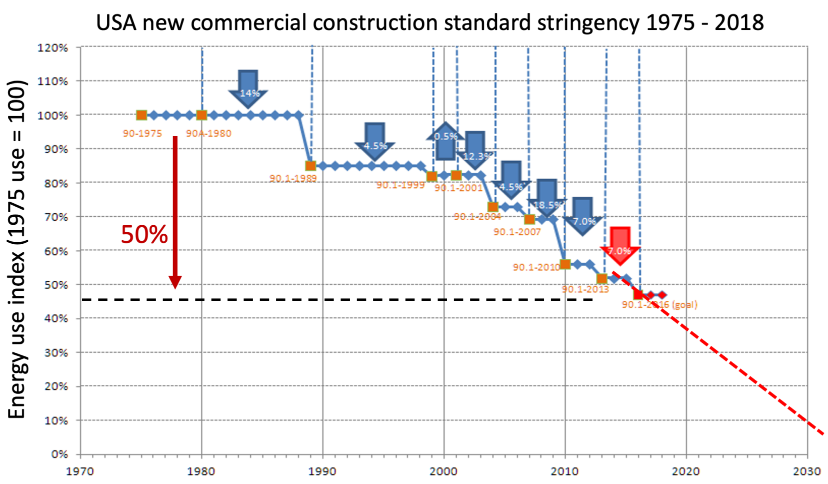

Figure 10: Effect of ASHRAE 90.1 on US building energy efficiency, 1970 - 2018 [32].

2.2 Equipment-Level Energy Efficiency and Control

Government agencies rate HVAC systems using data obtained at a small number of fixed thermodynamic conditions and fixed actuator settings, with the control function disabled [7]. This practice dates from the time when HVAC and refrigeration systems used fixed-speed fans and compressors, and fixed-orfice expansion devices such as capillary tubes, so that these measurements accurately represented expected real-world energy performance. Over time, however, manufacturers have introduced variable speed fans, compressors, and electronically actuated expansion valves, and increasingly complex control algorithms, so that the system can operate efficiently across a broad range of conditions – in principle. Unfortunately, in actual operation with the control function enabled, some variable-capacity systems may provide significantly lower energy efficiency compared to their ratings [**?**], particularly if little attention is paid to how systems interact within a building. This is sometimes caused by poorly behaved feedback control algorithms.

As a result, there is increasing interest in defining more realistic load-based testing methods to determine the energy ratings for equipment, and the California Energy Commission (CEC), the Canadian Standards Association (CSA) and Japanese rating agencies have all have drafted proposed standards that implement such methods [10]. For example, Figure 11 illustrates a proposed equipment rating test in which the compressor speed is varied continuously from high to low speed with the control function enabled. New rating tests are no simple matter, because of the wide variety of equipment, and also because a load-based test is considerably more difficult to execute compared to the existing standards, which apply a constant fluid state (temperature, humidity) at the equipment boundaries. Nonetheless, widespread adoption of new efficiency standards and measurement protocols that include the activation of control systems are inevitable.

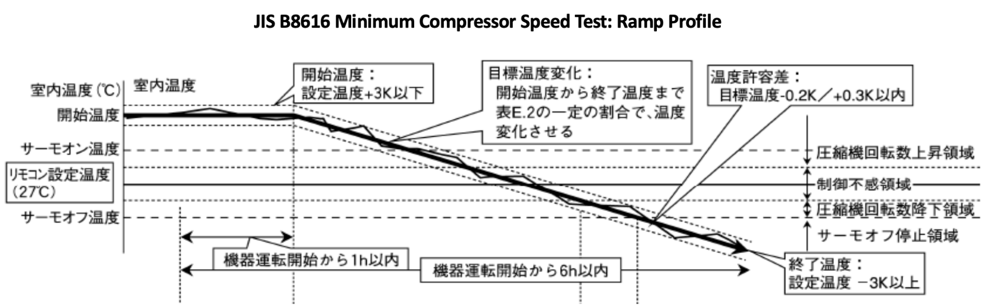

Figure 11: JIS B8616 Minimum Compressor Speed Test.

2.3 F-Gas Regulations

Fluorinated refrigerants (F-Gas) are potent greenhouse gasses, accounting for approximately 25% of global warming [47], with recent research showing that they are responsible for half of Arctic climate change in the past 50-60 years [37]. New government regulations are intended to reduce the amount of refrigerant that any one company may sell annually. For example, European F-Gas Regulation 517/2014 is an EU legislative instrument with provisions to reduce F-gas use as shown in Figure 12. Japan and China[3] have implemented, or are implementing, similar laws. This trend, as well as local building codes that incorporate ASHRAE Standard 15 [9], which limits the application of VRF products in some buildings to mitigate health and safety risks of refrigerant leaks, are pressuring all manufacturers to reengineer products to use less refrigerant and/or new refrigerants that have lower global warming potential (GWP).

In response to these trends, many equipment manufacturers have switched the refrigerant used in products from R410A to R32, which has a 3x lower GWP, but is mildly flammable. (Adoption of R32 lags in the US, largely because of the flammability risks.) In the near future, many of those

[3] http://www.env.go.jp/earth/ozone/hiyasu-waza/eng/revised_f-gas_law_in_japan.html

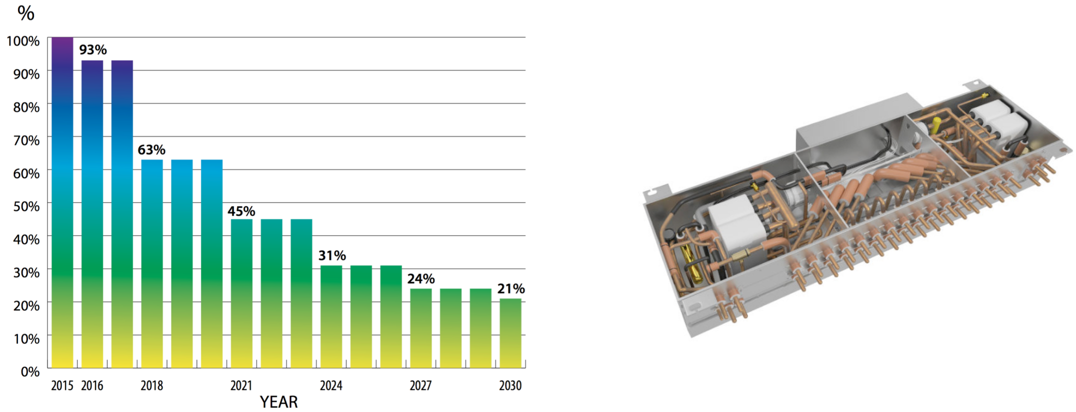

Figure 12: EU limits on F-Gas use in terms of GWP-equivalent of CO_2 (left), and a HVRF branch controller, showing heat exchangers, valves and water pumps (right).

still using R410A have committed to switching to R-454B, a zeotropic blend of R-32 and R-1234yf that has GWP that is 78% lower than R-410A.[4] (R-1234yf is replacing R-132 in the automotive industry, and has a GWP of less than 1.) This refrigerant presents challenges to engineering efforts because it has temperature glide, meaning the refrigerant temperature changes as a function of pressure in the two-phase region (See Figure 1), which makes physics-based models more complex, and may impact control design at the equipment level.

2.4 Decarbonization and Electrification

Buildings account for 40% of the primary energy consumption, 74% of the electricity consumption, and 39% of the CO2 emissions in the United States, with space heating and cooling being large contributors. In the US, over 70 U.S. cities have pledged to become "carbon neutral" by 2050. This is a potent trend, since building codes are governed by state and local municipalities in the U.S., rather than the federal government. Several municipalities already have outlawed new natural gas installations and have mandated or incentivized all-electric new construction for some types of buildings [44]. Similar trends are also apparent in Europe and Asia. Electric heat pumps offer the only viable alternative for space heating and cooling. In fact, a Siemens study, commissioned by the city of San Francisco, reports that market adoption of electric heat pumps is the single, most impactful action that the city may take to achieve its objective of an 80% reduction in emissions by 2050 [38].

2.5 High-Performance Buildings

High-performance buildings are those which exceed the minimum requirements defined by building codes and regulations, such as ASHRAE 90.1. Zero Energy Buildings (ZEB) in particular, defined as those which consume less energy than they produce on an annual basis, are made possible through advanced building envelope design that maximizes thermal insulation, minimizes undesired air infiltration, minimizes solar fenestration through shading and use of high performance glass, and may allow for natural ventilation when weather permits. Electric loads must be minimized by using natural daylighting and low energy lighting. Such buildings are often "all electric," using

[4]https://en.wikipedia.org/wiki/R-454B

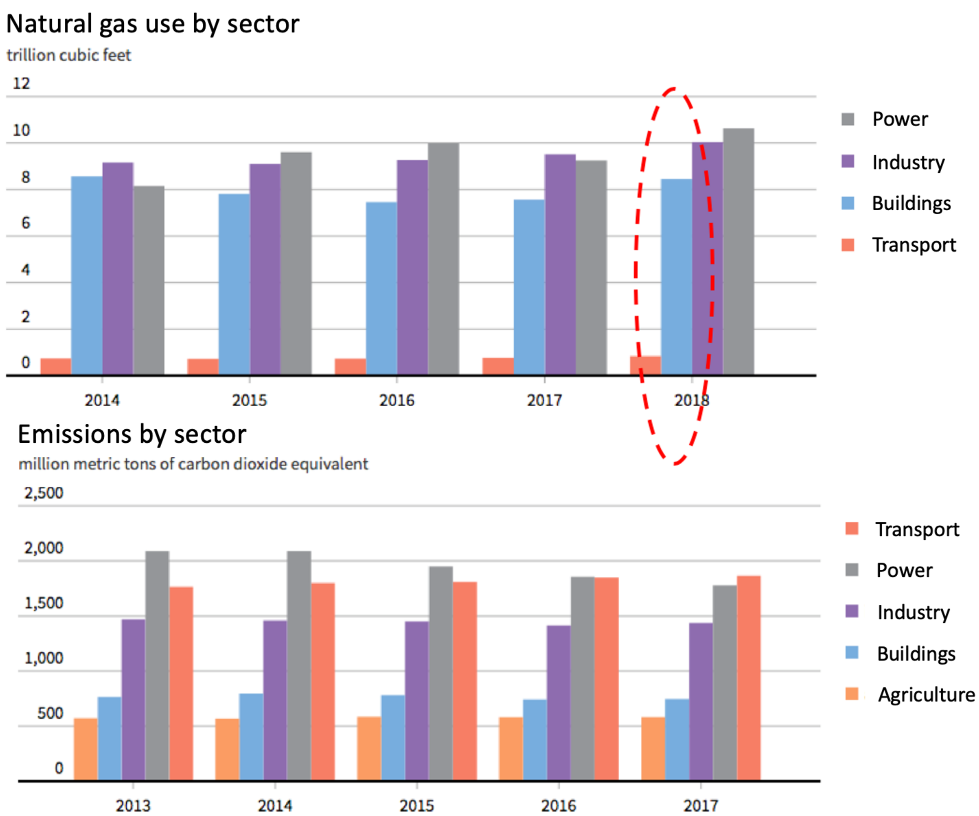

Figure 13: Natural gas use and emissions by sector, showing buildings to be a primary target.

heat pumps for HVAC and hot water, and exploiting waste heat where possible. With sufficiently low power loads, on-site photovoltaic generation can meet building needs on average. Significant attention must be paid to functional integration of the building electrical generation, supply, and distribution systems, HVAC, elevators, and lighting. From a control perspective, coordination and functional integration among subsystems, set-point and schedule optimization, and performance monitoring to make visible key aspects of building energy usage and identify and correct any problems are all critical.

Figure 14 shows Mitsubishi Electric's ZEB SUSTIE[5] building, a four-story office building located in Ofuna, Japan, completed in 2020. It was the first building in Japan to be certified as ZEB during construction, and during its first year of operation, achieved its ZEB objective, generating more power than it consumed. The building roof and south-facing window shades are covered with photovoltaics which generate sufficient power to meet demand from HVAC, lighting and elevators. The HVAC system is a highly modular all-electric VRF system, with optimized set-points and sequences. Hot water is generated using a CO2-based heat pump, and its "waste heat," (actually cold air) is used to help cool the building in summer. The BMS automatically opens some windows to provide natural ventilation when weather is favorable. All of the component technologies are available commercially.

[5] https://www.mitsubishielectric.com/en/about/rd/sustie/index.html

Figure 14: Mitsubishi Electric's SUSTIE building in Ofuna, Japan.

2.6 Renewable Electric Generation

On the energy production side, renewable sources such as wind and photovoltaic generate 20% of U.S. electricity, recently overtaking coal [4]. But their temporal variability poses a problem for stable grid operation, and this may limit growth. Building automation systems may offer at least a partial solution. A building has a certain degree of demand flexibility and can reduce or temporally shift some of its thermal energy use without negatively impacting occupant comfort. Demand response allows the grid operator to reduce electric load from buildings by resetting HVAC set points, providing a means to reduce peak demand, and is useful in extreme weather events. Traditional DR is deployed occasionally in response to price signals or explicit requests from the grid. Today, growing distributed energy resources (such as on-site solar PV and energy storage) and more frequent and extreme weather events are increasing the need and opportunity for integrating buildings into the grid. By taking advantage of building demand flexibility through advanced sensing, communication, and control technologies, it is possible to continuously optimize building operations in response to changing renewable supply levels, dynamic electricity prices, or other grid signals.

In 2019, the United States Department of Energy announced a Grid-interactive Efficient Buildings (GEB) Initiative, which aims to combine energy efficiency and demand flexibility with smart technologies to deliver greater benefits to building owners, occupants, and the electric grid [35]. A follow-on roadmap projected that national adoption of GEBs could yield US$100–200 billion in savings to the U.S. power system over the next two decades and 80 million tons annual reduction in CO2 emission reduction by 2030 [39].

3 Consequences and Implications

What are the technical and business consequences of these trends? Firstly, since the 1970s, energy efficiency initiatives have incentivized manufacturers to develop products with continuously variable actuators, replacing on-off components in order to improve efficiency at part load, enabled by the availability of inexpensive inverter technology. Compressors became variable speed. Electronic Expansion Valves (EEVs) replaced fixed-orfice capillary tubes. Variable speed fans and pumps replaced fixed speed fans and pumps. VAV systems replaced Constant Volume systems.

Of course, all these continuously variable actuators need to be controlled. The problem is, **at all levels, ranging from HVAC components up to the building level, these systems**

are increasingly dynamically coupled, interactive and multivariable. For example, the four-zone VRF system shown in Figure 5 has ten continuously variable actuators and 11 sensed variables. As a dynamical system, it is highly interactive from control inputs to measured outputs. Some VRF systems have hundreds of indoor units and dozens of outdoor units, ganged together.

Secondly, electrification means replacing fossil fuel-burning equipment used for heating (furnaces, boilers and such) with electric heat pumps (providing a lot of tail wind to the HVAC business). One challenge is to develop heat pumps that operate in more extreme climates. Heat pumps need to operate in colder outdoor air temperatures in winter, and higher ranges of outdoor temperatures in the summer. Moreover, again incentivized by energy efficiency, manufacturers have recognized the need to operate compressors over a wider range of frequencies. While today's heat pump might operate its compressor from about 25Hz to 125Hz, tomorrow's will operate from about 5Hz to perhaps 250Hz, giving a lower minimum capacity for very low load operation, and a higher maximum capacity for high load situations. The lower minimum capacity prevents start-stop operation at low loads, saving energy because the start-up transient draws higher than average amounts of power, while the higher capacity allows for application in increasingly warmer climates.

With the increased range of operation come constraints. The equipment must operate within a set of limits, or it can be damaged. There are upper and lower limits on several process variables, and these must be enforced by the control algorithm. In addition, there are limits to the ranges of actuation. Thus, the challenge here is nonlinearity: **The dynamics of a vapor compression system change as a function of operating condition, and constraints must be enforced in operation. This makes the control design problem nonlinear.** Gain scheduling and selector logic are often used, but as larger sets of operating conditions are considered, practical design methods may not scale.

Thirdly, HVAC systems are increasingly dynamically coupled with other building systems. Physically, they are coupled to the building's thermo-fluid dynamics. But HVAC systems are also integrated with other building systems through the system-level and building-level controls. For example, automatically operable windows provide natural ventilation, and this couples the envelope and outside weather to the HVAC system. Demand response and emerging grid-interactive technologies couple the electric supply to the HVAC system. As such, a buildings are slowly evolving toward complex systems-of-systems.

4 Industry Needs-Driven Innovation

In light of the trends and consequences, what are the industry's needs as it pertains to control? For one thing, precisely what the control community offers: Rigorous, mathematically-oriented model-based design and analysis. Many of today's control challenges can be met by applying mature control theory. However, there do remain some difficulties that, if addressed, will enable more robust products to be developed which, when properly integrated, can move ZEBs from demonstration to mainstream.

4.1 Next Generation Modeling & Analysis Tools

The foundations of all popular CAD tools (Matlab, TRNSYS, EnergyPlus, Modelica) were laid in the 1970s-1990s. Although all of these tools proved to be remarkably prescient, and each have evolved in the decades since, certain aspects are getting long-in-the-tooth. For example, Matlab Simulink is very popular and useful for simulating some types of relatively simple HVAC systems, but its signal-flow model of computation requires the user to predetermine causality of variables (the modeler must choose the states, and usually manually express their derivatives in terms of other states), and its numerical solvers simply do not stand up the needs of numerically stiff, nonlinear and large scale HVAC systems, especially those that include air-source heat exchangers. Modelica overcomes these limits, and is well suited for the purpose, but the analyses steps such as causalization and tearing are coded into the compilers with the specific intent to generate simulation code, so are largely hidden. It is not possible to access partial results, which might be useful in dynamic analysis. Resulting code is structured for use by state-of-the-art and highly optimized implicit numerical solvers, such as DASSL. However, these are largely single-threaded algorithms that use one processing core (although recent versions allow for some aspects of numerical solution calculation to be parallelized). Finally, the generated code is intended exclusively for simulation use on a desktop computer, requiring a modern operating system.

In terms of next generation platforms, some of the needs include the following.

1. **Symbolic Nonlinear Model Reduction.** Because (1) can be large scale and numerically stiff, it would be useful to reduce its dimension and also eliminate fast modes, akin to computing a singular perturbation [43, 29]. How do we compute compute a time-scale separation parameter ϵ as a function of physical parameters, as well as both fast and slow reduced-order models? A designer might want to compute multiple reductions for different values of ϵ, and conduct dynamic analysis on the fast and slow subsystems. This is a mixture of both symbolic calculus and numerical analysis. Model scale, including a large number of parameters θ makes this a challenge, although sparseness helps.

2. **Symbolic Jacobian calculation.** The Jacobian of (1) is central to application of robust multivariable control theory. Because the systems are large scale and stiff and the Jacobian is numerically ill-conditioned, it is important to have an exact, symbolic representation available. This is enabled by existing automatic differentiation (AD) technology, but is made challenging by the prevalence of highly nonlinear and nonsmooth fluid property functions in (1). Some Modelica tools can compute analytic Jacobians, but they require the user to provide some of the component-level derivatives i.e., the analytic Jacobian calculation is not fully automated.

3. **Parallelized implicit solvers.** Parallelizable algorithms for numerical calculation of solutions to (1), which can exploit multiple cores or GPUs, are needed. Existing algorithms such as DASSL are single-thread. Some aspects of the calculation of numerical solution, specifically the solution of the linear system of equations that lies at the heart of DASSL, can be parallelized, but often this provides little speed up. What else can be parallelized? Building simulation is an especially challenging application because of the stiffness and also long simulation times e.g., often one year of TMY weather data.

4. **Multi-Mode Systems.** Many HVAC systems are multi-mode, meaning subsets of states and equations may become active or inactive as a simulation evolves. For example, if an indoor unit is turned off, it may not be necessary to simulate its dynamics for some use cases. However, existing tools assume the same structure and number of states and equations for an entire simulation. New compiler algorithms are needed to transform acausal descriptions of multi-mode systems into simulation code that allows for a changing dimension of the state and number of equations at run time[11].

5. **Embedded Code Generation.** Modelica compilers generate code appropriate for desktop simulation. However, there is a need to take a Modelica representation of a controller and generate code meant for embedded application, which typically lacks operating system support. The developing eFMI standard aims to close this gap, targeting the automotive

industry initially, and there is significant opportunity to mature this technology and target it to the building automation industry.

The developing Modeling-Toolkit.jl [33] written in the Julia language, is a modeling package built on a symbolic computational algebra framework, and may provide a platform for future development. Like Modelica, it allows for users to construct large, acausal models from components, and it is equation - based and object oriented for organization. It uses similar algorithms for model analysis and causalization, such as Pantelides algorithm for DAE index reduction. However, Julia and the Modeling-Toolkit allow for a broader use of symbolic analyses, such as automatic differentiation, so that a Jacobian or Hessian of a model may be computed symbolically, ensuring that it is highly accurate when evaluated.

4.2 Robust Equipment-Level Control

HVAC equipment is increasingly multi-variable, interactive, and nonlinear, needing to operate over increasingly wider operating ranges. Although many feedback control loops are designed as SISO, the system as a whole needs to be analyzed as MIMO using modern multi-variable control methods such as disk margins [48, 40]. Often gain scheduling of feedback gains is used to ensure robust stability, but the dimension of scheduling variables can be large. For example, VRF system dynamics change as a function of indoor and outdoor temperature, heat load (or compressor speed), some fan speeds (which are not automatically controlled), and various geometric measurements. Gain scheduling on all of these is cumbersome. In addition, controllers contain selector and override logic to enforce prioritized constraints on process variables, and stability margins need to be guaranteed for all possible combinations of active constraints, in all operational conditions. For these systems, it is common that a constraint is active for long periods of time, and often these situations are considered normal operating modes. In other words, constraints are not to be *avoided*, but are actually designed into normal system operation. It is not uncommon to find an existing constraint-enforcement architecture cannot be extended to handle a new constraint as a design evolves or requirements change. In this case, the entire design might need to be revised, and there is no guarantee that a similar control architecture can be found. The process is time consuming, trial-and-error, and somewhat ad hoc. Finally, designs may be done at a few nominal conditions, but then need to be validated for hundreds to thousands of different conditions to ensure robust performance.

Model Predictive Control (MPC) holds some promise. As a design methodology for multivariable systems that can enforce input and output constraints, it seems well-suited to many HVAC equipment-level control problems. In most cases, output constraints can be considered “soft,” so that feasibility issues associated with real-time optimization are minimal. But several challenges must be addressed before MPC can succeed conventional selector-logic architectures. First and most importantly, MPC must address robustness with respect to model uncertainty. Robustness margins must be computable using models, so that robust stability can be guaranteed over large variation in the plant and operating conditions [13]. This may require introduction of nonlinear predictive models, so attention to efficient and robust nonlinear model reduction is required. Second, an MPC design needs to consider robust state estimation using available production sensors. This is often overlooked in the literature. Third, hybrid system issues such as start-up and shut-down must be addressed. How is an MPC started up, and modified when parts of an HVAC system are turned on and off? How does the designer deal with different modes of operation? These issues are considered directly by the designer in a selector-type architecture, where a start-up sequence typically closes one loop at a time. Forth, algorithms must be extremely computationally efficient for real world product use. Significant cost constraints limit the available computing power, requiring memory efficient and computationally efficient solver algorithms. These barriers currently limit MPC to laboratory experiments, but the pressures to develop a robust control methodology that replaces ad hoc selector logic methods will continue to grow. Each is a challenging research issue in its own right, requiring mathematical rigor to solve, but all require solution for MPC to have a broad application in this industry.

5 Vision Driven Innovation

What opportunities for vision-driven research might lead to meaningful innovation in this industry segment? Here we describe three opportunities: Digital Twins, Model Predictive Control and Grid-Interactive Buildings.

5.1 Digital Twin

A *digital twin* is a set of computer models that serve as a real-time digital counterpart of a physical object or process. The term was coined in the early 2000s in the context of Product Lifecycle Management (PLM) [27] to mean a set of computer representations of a product as it evolves through its lifecycle, from design to manufacture, then to operation, and finally to disposal. The digital twin was envisioned as an electronic repository of all aspects of design, such as 3-D CAD drawings and engineering simulation models, in addition to operational descriptions such as bills of process and operation. It is maintained throughout the product lifecycle via a real-time data stream of measurements obtained from the physical object. It is used to monitor and predict the behavior of the product in operation in its physical environment for diagnostics purposes, or in a variety of interrogative use cases in which future or past scenarios are analyzed to improve the design or operation of a product.

For our purposes, define an HVAC system digital twin narrowly to be a physics-based simulation model that is combined with measurements and used in real-time operation of the equipment. It may provide a range of benefits such as the following.

- **Virtual Sensing:** Heat flow through a heat exchanger, which is expensive to measure directly, especially for a direct expansion HEX, may be estimated using a model together with a limited set of temperature and actuator measurements. If sufficiently accurate, this may serve as a utility-grade meter for billing purposes.

- **Diagnostics**: The amount and location of the refrigerant charge inside HVAC equipment, which is difficult to measure directly, may be estimated and used to identify costly refrigerant leaks or conditions that cause refrigerant maldistribution, which can reduce energy efficiency and product reliability.

- **Model Predictive Control**: The digital twin model may be integrated into a product-level or building-level Model Predictive Control (MPC), which can command actuator values that optimize a cost function, such as energy use, over a time horizon, and also enforce constraints associated with the equipment or building operation.

For each of these use cases, a dynamic model of the HVAC equipment and possibly the building, is combined with real-time measurements in order to estimate a quantity of interest that is otherwise unavailable or difficult to measure directly. Since dynamic models are used extensively in product development, it seems reasonable to reuse them for this purpose. However, their use in the operational lifecycle phase differs significantly from their use in product development, and the dynamic model must be substantially modified, representing several research challenges.

1. **Estimation of States and Boundary Conditions**. From a control theoretic point of view, a digital twin is a state and boundary condition estimation problem: The model initial condition $x(t_0)$ and the disturbance (boundary condition) $d(t)$ must be estimated from feedback of available measurements $y(t)$ and the available model, which can then be used to predict future behavior. But there are challenges to be addressed. First, most state estimation schemes such as the Extended Kalman Filter (EKF) and its variants are formulated in discrete-time, and are structured as a recursive algorithm with a prediction step and a correction step. The correction step modifies the state prediction in order to assimilate measurements. But the model (1) includes states and relationships among states that are hard constraints. Most are implicit to the model and are not violated in typical simulations, given a set of physically consistent initial conditions, which is always the case when the model is used in a product development use case. For example, if humidity of the air is considered, there is a maximum limit to the amount of water that air can hold in the gas state at any given pressure and temperature, corresponding to 100% relative humidity. The model structure and physically consistent initial conditions prevent this constraint from being exceeded in a simulation use case. However, a state estimator might modify one or more states at an update step causing the relative humidity to exceed 100%. This causes the subsequent prediction step to fail numerically, because the model was simply not intended for this non-physical condition. Indeed, air leaving a cooling coil is usually very near 100% humidity, so this particular example is hardly a corner case. Flow direction is another difficult constraint to handle. The model (1a) makes no a priori assumption on flow direction among any of its control volumes (which range from refrigerant tubes to rooms). Flow direction is a computed variable, and the structure of (1a) changes as a function of flow direction (downstream fluid states depend on upstream ones). If the update step changes control volume pressures in a way that changes the flow direction, then the model used in the update step is inconsistent with the model used in the prediction step. Again, this is far from a corner case, because pressure drops between control volumes can be very small in normal operation of an HVAC system or in a model of building airflow.

   Therefore, the state and boundary condition estimation problem is constrained, with many hard constraints on the state, boundary conditions and parameters that need to be made explicit, although most are not expressed as *explicit* constraints in simulation use cases. Constrained estimation algorithms such as the constrained Ensemble Kalman Filter (EnKF) [8, 19] or optimization-based constrained estimators [25] need to be considered, but even then the designer faces severe challenges. The model (1) is large scale, nonlinear and numerically stiff, so that Jacobians are numerically ill-conditioned. This causes estimation algorithms such as an Extended Kalman Filter to have small domains of convergence, making them non-robust in practice. Data Assimilation methods used in weather forecasting may offer an approach. These problems are also large scale, nonlinear, and typically include diverse measurements. However, weather models are usually not numerically stiff and do not need to include phase change, compressible fluid flow or widely varying fluid properties to the extent that HVAC models require.

2. **Model Calibration**. Values for elements of the parameter vector θ, including building material parameters and geometry, along with the configuration of the HVAC system, need

to be calibrated for every instance. If we consider a digital twin of an entire building and HVAC system, then the number and diversity of parameters can be overwhelming and require a large amount of manual labor, although digital building representations (BIM) are evolving and in principle can contain much of the needed data. Even if we consider a digital twin of a factory-built HVAC unitary unit, differences in installation details imply that parameters will vary by installation and therefore require calibration.

For digital twins to move beyond one-off demonstrations, calibration will require robust automation, not only to calibrate fixed parameters, but also to configure the model structure. Moreover, some parameters can vary over the lifecycle of a product, and estimating values for these is valuable in order to ascertain the health of the equipment. For example, model (1a) is relatively sensitive to the heat transfer coefficients on the air side of a HEX, and these can change over time as the HEX accumulates dirt or corrodes. These may be considered constant for a product development use case, but for a digital twin, an historical evolution is of interest. An important problem is to estimate the amount of refrigerant in a vapor compression cycle. This may be assumed to be constant in a desktop simulation, but slow leaks are common in the field and it is of interest to multiple parties to identify the leak rate and location to facilitate repair. How should an estimator be designed to estimate this slowly - varying parameter? Such a problem is deceptively difficult because of the enormous time scales involved, and considering that the refrigerant charge is not a conserved quantity, i.e., neither the numerical algorithm used for integration, nor the model itself, are constructed to conserve charge, so that it might "drift," or change very quickly, during a long-term numerical simulation. The solution to this problem will require reformulation of the model (1a), in addition to new types of parameter and state estimation algorithms, and new methods of numerical integration.

3. **Tools and Platforms**. CAD tools used to represent and simulate the model (1) are intended for desktop simulation in a product development use case. They do not easily allow for integration of real-time information from measurements, or scaling up of models beyond a single simulation, although most support parametric simulation studies. New tools and platforms need to be developed that will enable development of digital twins beyond desktop simulation, and support eventual deployment. One promising technology for this is the Functional Mockup Interface (FMI)[6], which is a standard for sharing and simulating models created in Modelica. A Modelica model may be compiled into a Functional Mockup Unit (FMU), which is an executable software package that allows the simulation to be executed on a wide variety of platforms, such as in Python, Matlab or Microsoft Excel. Its primary intent is for model sharing, but an FMU allows for two operations that enable realization of constrained estimators. One important feature of the FMI standard is that it depends on only one other standard: ANSI C. This is stable, so that an FMU can be expected to have a long lifetime. This is an important consideration when developing digital twins for HVAC equipment in buildings, which have lifespans of approximately 30 years.

5.2 Building-Level Model Predictive Control

For building-level control, MPC has gained a lot of attention from the research community for more than a decade [20]. At this level of the control hierarchy, economic MPC is technically well-suited to minimization of energy or power consumption or a related cost by recursively solving a real-time optimization problem, which is typically parameterized by HVAC system set points,

[6]https://fmi-standard.org/

subject to a predictive model and constraints associated with human comfort. Many building HVAC MPC strategies have been reported to achieve energy efficiency goals [31, 36] or incorporate demand response objectives [16, 15, 22, 45], with benefits of energy savings from 15% to 50%. In contrast to the equipment-level, MPC at this level is less concerned with closed-loop stability and robustness, and more concerned with optimization of a economically meaningful objective function, usually over a longer time horizon of hours to days [13]. (On the other hand, instabilities and oscillations have been observed at this level too [20].) The main benefit of MPC is that it is a systematic methodology for multivariable optimization of energy consumption (or similar economic cost) subject to constraints, which can be non-obvious because of multivariable interactions among systems and subsystems.

Yet, despite its promise and the research efforts to date, transferring MPC into industrial practice remains in its early stages, due to its cumbersome and time-consuming design and implementation procedures (especially building and calibrating predictive models, and tuning of controller parameters) and online computational requirements, etc. [18]. Automating the process of model construction, calibration, and integration with tools that provide for real-time measurements and numerical optimization is needed before MPC at the building level can move beyond demonstration. As it stands, many demonstrations are accomplished with highly skilled but relatively low cost, subsidized student labor. This environment allows for models to be constructed, calibrated, interfaced to data streams and optimizing software for the purpose of conducing a short-term demonstration. But even if this demonstration shows considerable performance improvement relative to some baseline, the process of creating and maintaining the demonstration is not sustainable nor does it scale beyond the demonstration. If MPC is to be adopted as a building-level methodology, the costs and skill level associated with its deployment must be no greater than what exists today, implying it must be highly automated and result in a robust and transparent (easy to understand, install, debug etc.) technology. How to do this is a serious intellectual research question.

Many of the barriers to adoption of MPC at the building level are similar to those of a digital twin, which can be considered the predictive model in an MPC algorithm, but the business interests in the digital twin are better aligned. Businesses as end users of HVAC equipment (both building owners and tenants) increasingly need to monitor and account for their energy consumption and CO2 footprint, for sustainability reasons. Monitoring of energy consumption and continuous improvement to achieve corporate sustainability goals is rapidly becoming a business practice norm, much like financial accounting. This will motivate these parties to invest in development of digital twins. At the same time, equipment manufacturers have strong interest in understanding how their products behave in operation for several reasons: To close the loop in development and improve future products, to deepen relationships with customers, and possibly to develop revenue steam services. As such, multiple parties have aligned incentives to develop digital twin technology, if only for purposes of performance monitoring and off line scenario planning. As it develops, using the digital twin as the core of an automated economic MPC becomes more feasible. Therefore, it is likely that the digital twin will proceed industrial application of MPC at the building level.

One potentially interesting and unexplored research topic is application of MPC over considerably shorter time horizons than are usually considered at the building level. Typically economic MPC time horizons span from many hours to many days, or even longer. However, with solar PV supplying power to a building-scale microgrid, it is possible to consider predictive control over a horizon of minutes, and use video sensors to monitor the sky and predict the PV power supply over a few minute time horizon, which can vary on partly cloudy days. An MPC at this time scale could manipulate the HVAC system (and other building systems such as hot water and lighting) such that its peak (and average) power consumption remains below the available PV supply, which would reduce or eliminate short time-scale demand from the grid. This would require integration of the HVAC equipment control, possibly allowing the building-level control to manipulate constraints in the equipment, such as the maximum compressor speed, which then enforces these constraints locally, instead of manipulating zone set points. Since the variation in solar load is on a relatively high frequency, the average zone temperature would be minimally impacted, so human comfort would be maintained.

5.3 Grid-Interactive Buildings

Beginning in 2019, the U.S. Department of Energy released a series of reports about grid-interactive efficient buildings (GEBs) that use smart technologies and on-site demand side resources (DSRs) to provide demand flexibility while co-optimizing for energy cost, grid services, and occupant needs and preferences, in a continuous and integrated way [35, 39]. GEBs can provide demand flexibility via five modes: efficiency, shedding, shifting, modulating, and generation as shown in Figure 15.

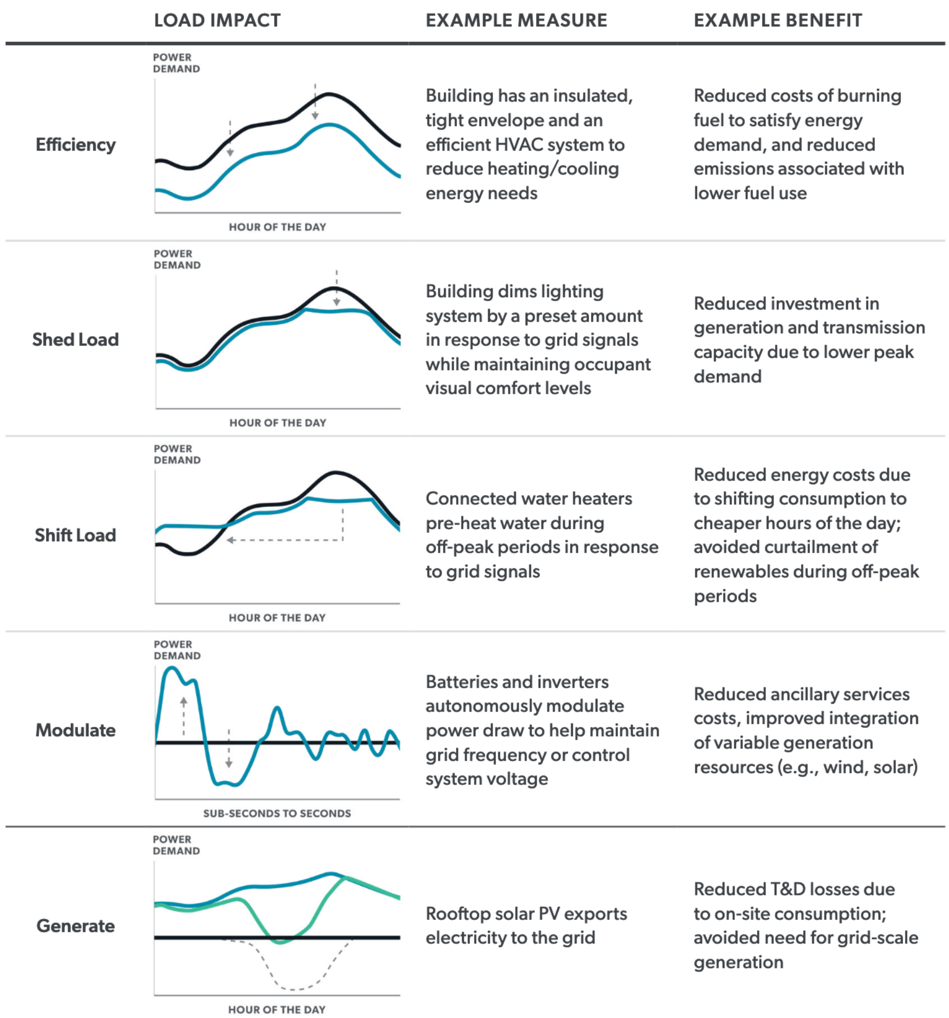

Figure 15: Five modes of electric demand flexibility [39].

Energy efficiency and traditional demand response (load shifting) are decades-old strategies intended to reduce the overall energy demand from the grid and curtail energy use during peak demands or emergencies. While these strategies are typically employed separately, to address the rising needs for effective demand-side management, buildings must be designed and operated to provide enhanced benefits for buildings' owners, occupants, and the electric grid. Grid-interactive efficient buildings (GEBs) are receiving increased attention in this context for their ability to reduce building energy use and provide grid services through demand flexibility. (A good example of aligned incentives.). The rapid advancement of building technology capabilities and the increasing deployment of behind-the-meter distributed energy resources such as PV provide greater opportunities for buildings to contribute to managing occupant satisfaction and the grid while decreasing the overall cost of building ownership.

Currently, building management operations focus on energy efficiency, cost savings and occupant comfort. Legacy systems do not provide any grid services. Different grid services require varied levels and types of flexibility. For example, balancing generation capacity requires longer load shifts while balancing the grid frequency benefits from fast and short-term changes to usage. The ability to provide more frequent holistic slower services and fast-acting services without compromising building functions differentiates GEBs from what is being done today. Of course these strategies all require system-level control, which must provide robustly stable performance over wide range of operating conditions. This represents a considerable research opportunity to the controls community, as it has received very little attention as an issue in the development of GEBs.

Buildings can contribute to slower-acting services (e.g., generation capacity relief, contingency reserve) by both lowering the overall need for generation through energy efficiency as well as reducing electricity use for a period (load shed) or changing the timing of electricity use (load shift). The fast-acting services (e.g., frequency regulation) are deployed to correct short-term imbalances in the grid, usually on time scales of seconds to minutes. It requires continuous modulation of building technologies such as battery inverters and variable frequency drives in fans and chillers in response to a signal from the grid operator. Qualified resources must respond to dispatched control signals (within seconds) to increase or decrease their electrical load according to the power grid's needs.

5.3.1 Demand Flexibility

Building demand flexibility must be provided within the limits dictated by the building systems and operation. The building demand flexibility is the acceptable modifications to the baseline electricity consumption (via load shed, shift, and modulation) with no undesirable reduction in the building primary functions. Potential methods and metrics for quantifying and predicting building HVAC demand flexibility have been proposed e.g., [30, 5, 41]. Practical methods, in the context GEBs, will need to consider the operation of on-site DERs and account for uncertainty sources affecting building operations. The dynamic aspect of demand flexibility is also vital as building operational efficiency and energy usage vary over time depending on weather conditions, occupancy needs, etc. As systems become interconnected and controlled, closed-loop robustness and stability margins for these strategies must defined and used for their design and implementation.

5.3.2 Control and Coordination Strategies

Broader market adoption of demand flexibility will benefit from scalable control and coordination strategies that can collaboratively engage and aggregate building energy efficiency and demand flexibility technologies, including the enhanced flexibilities offered through on-site thermal or electrical energy storage. While there are several research activities in this direction, a winning solution must be cost-effective, scalable, and easy to deploy with verifiable and sustained benefits. The potential control approaches can be broadly classified into heuristic or rule-based control (RBC) and optimization-based control such as MPC. Although RBC sequences are easier to implement, verify and understand, it is difficult for them to adapt to changing external conditions (e.g., variation in grid peaks) or take anticipatory actions, which are critical to realizing the full benefits of demand flexibility. On the other hand, MPC operates and plans over a time horizon and can optimize multiple objectives. MPC has shown substantial benefits (over RBC) in reducing building operation costs and activating energy flexibility and handling constraints. However, the approach requires system models, forecasting, and non-trivial computational complexities, as previously discussed. AI-based techniques, such as reinforcement learning, which learn over time by interacting with the environment have some potential. Still, this technique is relatively immature, doesn't scale well, and hasn't proven effective in real-world building applications. Therefore, there is a clear need for a unified control approach that agglomerates the advantages of MPC and learning-based adaptive techniques with the relative implementation ease and interpretability of rule-based controls.

5.3.3 Cyber-Security in Buildings

Modern Building Automation Systems (BASs), require connectivity among systems within the building as well as with outside entities, such as the cloud, to enable low-cost remote management, optimized automation via outsourced cloud analytics, and building-grid integration. As BMSs evolve towards open communication technologies, providing access to BASs through the building's intranet, or even remotely through the Internet, has become a common practice. However, increased connectivity and accessibility come with increased cyber security threats. BASs were historically developed as closed environments with limited cyber-security considerations. As a result, BASs in many buildings are vulnerable to cyber-attacks that may cause adverse consequences, such as occupant discomfort, excessive energy usage, or unexpected equipment downtime.

The rising demand for enhancing BAS cyber-security calls for a comprehensive understanding of the BAS cyber landscape. Recent research and studies have been focused on cyber-physical security on BASs, which mainly cover cyber-attacks, detection, and defense related topics. But few publications have focused on cyber-secure resilient control strategies specifically for BASs in commercial buildings. Generally speaking, in contrast to other domains that recently received substantial attention such as industrial control and automation systems [26], the security of BASs has been discussed in a less structured manner. An in-depth analysis is still needed to systemically address the cyber-security issues of BASs in the context of the emerging openness and connectivity of intelligent buildings. There are needs for:

- Developing cyber analytics solutions that can minimize the frequency of detection false alarms and accurately diagnose and localize cyber-attacks. Preventative strategies are needed as early alarms to catch cyber-attacks before they happen on BASs. Solutions that can differentiate cyber-attacks from physical faults are also needed to assure targeted response and fast recovery from the effects of adversarial events.

- Developing resilient strategies that can handle multiple simultaneous cyber-attacks and physical faults. Most studies focused on only one type of event at a time. However, multiple cyber-attacks and physical faults can occur simultaneously. Therefore, an attractive future direction is developing a flexible detection/defense/control solution to tackle diverse and concurrent cyber threats and faults.

6 Conclusions

So, what are the control research challenges that, if addressed, will enable and drive meaningful innovation in building automation in the 21st century? Control as a subject (a body of knowledge) is relatively unique in its recognition of robust stability and performance, and of the fundamental and rigorous trade-off between robustness and performance. HVAC systems in particular have slowly evolved from being relatively static systems where closed-loop stability was simply not an issue, into interactive, multivariable, hybrid and nonlinear dynamical systems where robust closed-loop stability is central to correct operation at all levels. This evolution has been driven by a need to improve energy efficiency, and has occurred so gradually over the past 50 years, that the issues of robust stability and performance is still not widely recognized as being critical. (Much like the boiling frog myth, but this is no myth.). From a needs point-of-view, building automation thus offers innumerable opportunities to those knowledgable in control. These problem-driven issues will only increase in their importance as buildings become more and more electric, dynamically interactive with the electrical grid, and the grid itself becomes more and more decentralized. Robustness needs to be emphasized and researched along side optimization of performance. Directly addressing these problems for real-world products, without changing the assumptions to suit the researcher's constraints or needs, will lead to improved energy and comfort performance of building automation systems, and also a deeper understanding of the gaps and needs in theory. None of these applications involve simply applying existing, well-known methods and tools.

In terms of vision-driven innovation, we have highlighted three areas: Digital Twins, Building-Level MPC, and Grid-Interactive Buildings. These areas represent opportunities for proactive researchers with new ideas, and are likely to remain areas of innovation for many decades to come. Combining real-time physics-based models with new algorithms that are needed to estimate unmeasured quantities, some of which are physical and others less so, offer opportunities to those interested in estimation and statistical learning theories. Building level MPC has been researched but has yet to find a strong industrial need, demands a high degree of expertise, and remains largely practiced as case studies. Yet it may offer a means to integrate physical systems (HVAC, grid for example) as well as models of computation (physics-based models with occupancy and behavior models for example, which are learned). There is opportunity here, but issues related to robustness, often neglected, need to be front and center. Grid-Interaction, and on-site generation, will grow in importance and will increasingly require some attention to dynamic interaction. Existing demand response algorithms, which reduce set-points for a specified period of time, will evolve, and the control field should play a role in shaping this evolution to ensure issues related to dynamic interaction are considered.

Finally, new types of collaborations among partners in academia, industry and government laboratories need to be imagined. Many of these challenges and opportunities are multi-disciplinary and non-trivial, and are simply not addressable in any meaningful way via the conventional student-constrained, short-term research paradigm common to academia. There is insufficient time for a graduate student to master the required fields and then make a contribution. Longer term research programs need to be established, with involvement from leaders in industry who are willing to share their knowledge, experience, problems, needs and challenges, which are often considered proprietary and kept secret, or are poorly understood. Incentives need to be established in academia and at government laboratories for impacting commercial technology and/or industrial practice. (This is especially so for engineering). On the other hand, industry needs to come together and communicate longer-term needs, opportunities and challenges, and then actively work together with partners in government and academia to gradually solve them. This is not today's funding model. Innovation in this industry evolves slowly, but this does not diminish its importance or compelling nature, and in fact demands new ways of collaborating.

References

[1] Buildings and climate change: Summary for decision - makers. UNEP's Sustainable Buildings and Climate Initiative (SBCI), 2009.

[2] *2020 ASHRAE Handbook: HVAC Applications*. ASHRAE, 2019.

[3] *2020 ASHRAE Handbook: HVAC Systems and Equipment*. ASHRAE, 2020.

[4] U.S. Energy Information Administration monthly energy review, May 2020.

[5] V. A. Adetola, F. Lin, and H. M. Reeve. Building flexibility estimation and control for grid ancillary services. In *High Performance Buildings Conference*, West Lafayette, IN, 2018. Purdue Herrick Labs.

[6] Azkia Afroz, GM Shafiullah, Tania Urmee, and Gary Higgins. Modeling techniques used in building HVAC control systems: A review. *Renewable and Sustainable Energy Reviews*, 2017.

[7] Air-Conditioning, Heating & Refrigeration Institute (AHRI), 2111 Wilson Blvd, Suite 500. *Standard for Performance Rating of Unitary Air-conditioning and Air-source Heat Pump Equipment*, 2017.

[8] David J. Albers, Paul-Adrien Blancquart, Matthew E. Levine, Elnaz Esmaeilzadeh Seylabi, and Andrew Stuart. Ensemble Kalman methods with constraints. *Inverse Problems*, 35(9), Aug. 2019.

[9] Refrigerating American Society of Heating and Air-Conditioning Engineers. ASHRAE standard 15-2019 safety standard for refrigeration systems, 2019.

[10] Canadian Standards Association. Load-based and climate-specific testing and rating procedures for heat pumps and air conditioners.

[11] Albert Benveniste, Benoit Caillard, and Mathias Malandain. The mathematical foundations of physical systems modeling languages. Project-Team Hycomes 9334, INRIA, April 2020.

[12] Scott A. Bortoff, Dan J. Burns, Chris R. Laughman, Hongtao Qiao, Claus Danielson, Abraham Goldsmith, and Stefano Di Cairano. Power optimizing control of multi-zone heat pumps. In *IEEE Conference on Control Technology and Applications*, pages 826–833, Aug. 2018.

[13] Scott A. Bortoff, Paul Schwerdtner, Claus Danielson, Stefano Di Cairano, and Daniel J. Burns. H-infinity loop-shaped model predictive control with heat pump application. *IEEE Transactions on Control Systems Technology*, 30(5), Sept. 2022.

[14] Kathryn E. Brenan, Stephen L. Cambell, and Linda R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. SIAM, 1996.

[15] Sergio Bruno, Giovanni Giannoccaro, and Massimo La Scala. A demand response implementation in tertiary buildings through model predictive control. *IEEE Transactions on Industry Applications*, 55(6):7052–7061, 2019.

[16] JA Candanedo, VR Dehkordi, and M Stylianou. Model-based predictive control of an ice storage device in a building cooling system. *Applied Energy*, 111:1032–1045, 2013.

[17] Francois E. Cellier. *Continuous System Simulation*. Springer, 2006.

[18] Jiří Cígler, Dimitrios Gyalistras, Jan Široky, V Tiet, and Lukaš Ferkl. Beyond theory: the challenge of implementing model predictive control in buildings. In *Proceedings of 11th Rehva world congress, Clima*, volume 250, 2013.

[19] Vedang Deshpande, Christopher R. Laughman, Yinbo Ma, and Chris Rackauckas. Constrained smoothers for state estimation of vapor compression cycles. In *Proceedings of the American Control Conference*, 2022.

[20] Ján Drgoňa, Javier Arroyo, Iago Cupeiro Figueroa, David Blum, Krzysztof Arendt, Donghun Kim, Enric Perarnau Ollé, Juraj Oravec, Michael Wetter, Draguna L Vrabie, et al. All you need to know about model predictive control for buildings. *Annual Reviews in Control*, 2020.

[21] Hilding Elmqvist, Martin Otter, and Sven Erik Mattsson. Fundamentals of synchronous control in Modelica. In *Proceedings of the 11th International Modelica Conference*, pages 15–25, 2011.

[22] Luca Fabietti, Tomasz T Gorecki, Faran A Qureshi, Altuğ Bitlislioğlu, Ioannis Lymperopoulos, and Colin N Jones. Experimental implementation of frequency regulation services using commercial buildings. *IEEE Transactions on Smart Grid*, 9(3):1657–1666, 2016.

[23] Peter Fritzon. *Principles of Object Oriented Modeling and Simulation with Modelica 3.3: A Cyber-Physical Approach*. Wiley, 2015.

[24] Peter Fritzson, Adrian Pop, Karim Abdelhak, Adeel Asghar, Bernhard Bachmann, Willi Braun, Daniel Bouskela, Robert Braun, Lena Buffoni, Francesco Casella, Rodrigo Castro, Rudiger Franke, Dag Fritzson, Mahder Gebremedhin, Andraes Heuermann, Bernt Lie, Alachew Mengist, Lars Mikelsons, Kannan Moudgayla, Lennart Ochel, Arunkumar Palanisamy, Vitalij Ruge, Wladimir Schamai, Martin Sjolund, Berhard Thiele, John Tinnerholm, and Per Ostlund. The openmodelica integrated environment for modeling, simulation and model-based development. *Modeling, Identification and Control*, 41(4):1890–1328, 2020.

[25] Graham C. Goodwin, María M. Seron, and José A. De Doná. *Constrained Control and Estimation: An Optimization Approach*. Springer, 2005.

[26] Vitor Graveto, Tiago Cruz, and Paulo Simöes. Security of building automation and control systems: Survey and future research directions. *Computers & Security*, 112:102527, 2022.

[27] Michael Grieves. Origins of the digital twin concept. Technical report, ResearchGate, Aug. 2016.

[28] David Holmberg. Demand response and standards. *ASHRAE Journal*, BACnet Today & The Smart Grid Supplement:B23–B28, November 2011.

[29] Hassan K. Khalil. *Nonlinear Systems: Third Edition*. Prentice Hall, 2002.

[30] Han Li, Zhe Wang, Tianzhen Hong, and Mary Ann Piette. Energy flexibility of residential buildings: A systematic review of characterization and quantification methods and applications. *Advances in Applied Energy*, 3:100054, 2021.

[31] Wei Liang, Rebecca Quinte, Xiaobao Jia, and Jian-Qiao Sun. Mpc control for improving energy efficiency of a building air handler for multi-zone vavs. *Building and Environment*, 92:256–268, 2015.

[32] Bing Liu, Michael Rosenberg, and Rahul Athalye. National impact of ansi/ashrae/ies standard 90.1-2016. In *2018 Building Performance Conference and SimBuild*. ASHRAE and IBPSA-USA, September 2018.

[33] Yingbo Ma, Ranjan Anantharaman Shashi Gowda, Chris Laughman, Viral Shah, and Chris Rackauckas. ModelingToolkit: A composable graph transformation system for equation-based modeling composable graph transformation system for equation-based modeling. Technical report, Julia Computing, 2021.

[34] Modelica Association, https://www.Modelica.org/. *Modelica Language Specification Version 3.5*, Feb. 2021.

[35] Monica Neukomm, Valerie Nubbe, and Robert Fares. Grid-interactive efficient buildings technical report series: Overview of research challenges and gaps. Technical report, US Department of Energy, 2019.

[36] Frauke Oldewurtel, Dimitrios Gyalistras, Markus Gwerder, Colin Jones, Alessandra Parisio, Vanessa Stauch, Beat Lehmann, and Manfred Morari. Increasing energy efficiency in building climate control using weather forecasts and model predictive control. In *Clima-RHEVA World Congress*, number CONF, 2010.

[37] L. M. Polvani, M. Previdi, M. R. England, and K. L. Smith. Substantial twentieth-century Arctic warming caused by ozone-depleting substances. *Nature and Climate Change*, January 2020.

[38] Dennis Rodriguez. Reaching 80x50: Technology pathways to a sustainable future. Technical report, Siemens, 2020.

[39] Andrew Satchwell, Mary Ann Piette, Aditya Khandekar, Jessica Granderson, Natalie Mims Frick, Ryan Hledik, Ahmad Faruqui, Long Lam, Stephanie Ross, Jesse Cohen, Kitty Wang, Daniela Urigwe, Dan Delurey, Monica Neukomm, and David Nemtzow. A national roadmap for grid-interactive efficient buildings. Technical report, Lawrence Berkeley National Laboratory; The Brattle Group; Energy Solutions; Wedgemere Group; U.S. Dept. of Energy, 2021.

[40] Peter Seiler, Andrew Packard, and Pascal Gahinet. An introduction to disk margins. *IEEE Control Systems Magazine*, pages 78–95, Oct. 2020.

[41] Hong Tang and Shengwei Wang. Energy flexibility quantification of grid-responsive buildings: Energy flexibility index and assessment of their effectiveness for applications. *Energy*, 221:119756, 2021.

[42] Bernhard Thiele, Thomas Beutlich, Volker Waurich, Martin Sjölund, and Tobias Belmann. Towards a standard-comform, platform-generic and feature-rich Modelica device drivers library. In *Proceedings of the 12th International Modelica Conference*, pages 713–723, 2017.

[43] Mathukualli Vidyasagar. *Nonlinear Systems Analysis: Second Edition*. Prentice-Hall, 1993.

[44] Valerie Volcovici and Nichola Groom. The next target in the climate-change debate: Your gas stove. *Reuters*, 2019.

[45] Evangelos Vrettos, Emre C Kara, Jason MacDonald, Göran Andersson, and Duncan S Callaway. Experimental demonstration of frequency regulation by commercial buildings—part i: Modeling and hierarchical control design. *IEEE Transactions on Smart Grid*, 9(4):3213–3223, 2016.

[46] Michael Wetter, Wangda Zuo, Thierry S. Nouidui, and Xiufeng Pang. Modelica buildings library. *Journal of Building Performance Simulation*, 7(4):253–270, 2014.

[47] Donald J. Wuebbles. The role of refrigerants in climate change. *International Journal of Refrigeration*, 1994.

[48] Kemin Zhou, John Doyle, and Keith Glover. *Robust and Optimal Control*. Prentice Hall, 1996.