Quantitative Comparison of Data-Driven and Physics-Based Models for Commercial Building HVAC Systems

Datong P. Zhou*, Qie Hu*, and Claire J. Tomlin

***Abstract*— Commercial buildings are responsible for a large fraction of energy consumption in developed countries, and therefore are targets of energy efficiency programs. Motivated by the large inherent thermal inertia of buildings, the power consumption can be flexibly scheduled without compromising occupant comfort. This temporal flexibility offers opportunities for energy savings and the provision of frequency regulation to support grid stability. To realize these goals, it is of prime importance to identify a realistic model for the temperature dynamics of a building. In this paper, we identify a low-dimensional data-driven model and a high-dimensional physics-based model for the same system at different spatial granularities and temporal seasons using experimental data collected from an entire floor of an office building on the University of California, Berkeley campus. We perform a quantitative comparison in terms of estimates of the inherent thermal gains due to occupancy, open-loop prediction accuracies, and closed-loop control schemes. We conclude that data-driven models could serve as a substitution for highly complex physics-based models with an insignificant loss of prediction accuracy for many applications.**

I. Introduction

According to [1], residential and commercial buildings account for up to 40% of the total electricity consumption in developed countries, with an upward trend. Heating, ventilation and air-conditioning (HVAC) systems are a major source of this consumption [2]. Nevertheless, their power consumption can be flexibly scheduled without compromising occupant comfort, due to the thermal capacity of buildings. As a result, HVAC systems have become the focal point of research, with the goal of utilizing this source of consumption flexibility. From the point of view of energy efficiency, researchers have studied optimization of building control in order to minimize power consumption [3], [4]. Further, by participating in the regulation of electricity's frequency, buildings can assist in supporting the supply quality of electricity and the grid stability [5], [6], [7], [8].

All of the above research activities are based on a valid mathematical model describing the thermal behavior of buildings. Traditionally, buildings have been modeled with high-dimensional physics-based models such as resistance-capacitance (RC) models [9], [10], TRNSYS [11], and EnergyPlus [12]. These models are motivated by the thermodynamics of the building and explicitly model the heat transfer between building components. The advantage of such models is their high granularity of temperature modeling, but a drawback is their high dimensionality, rendering them computationally expensive. A large body of this work focuses on linear models, whereas physics-based models for commercial buildings with a variable air volume (VAV) HVAC system are bilinear in nature. Furthermore, existing model reduction techniques often result in a loss of interpretability of states [13] and disproportionate increase in the model's prediction error [14].

Motivated by these shortcomings, a new direction of research attempts to identify low-dimensional, data-driven models with Input-Output models [7] and semiparametric regression [15]. The intention is to alleviate the computational complexity in expense for coarser and less accurate temperature predictions.

The contribution of this paper is two-fold. First, we aim to improve existing data-driven model identification techniques. Unlike [16], [17], who model the evolution of the building's energy consumption without a specific control input, we identify a model for temperature evolution in multiple building zones amenable to control design, i.e. with airflows as inputs. Our model also differs from [15], which uses HVAC supply air temperature as the single control input, resulting in a simpler identification problem, but, on the other hand, offers less flexibility for control.

Second, and more importantly, we perform a *quantitative comparison* of data-driven and physics-based models in terms of open-loop prediction accuracy and closed-loop control strategies, based on the *same testbed* (the entire floor of an office building) using *experimental data* collected from the building, as opposed to simulated data. We conclude that a low-dimensional data-driven model is suitable for building control applications, such as frequency regulation, due to its minor loss of prediction accuracy compared to high-dimensional physics-based models, but significant gain in computational ease. To the best of our knowledge, the extant body of literature has analyzed data-driven and physical models for the identification of temperature evolution in commercial buildings only in an isolated fashion (in particular not on the same testbed), [3], [18]. In addition, some of these models were identified for fictitious buildings with synthetic data [19], [20], while others used experimental data collected under environments with little or no disturbance, e.g. without occupants [7]. Our work differs from these existing works in two aspects. First, we use experimental data to identify models for a multi-zone commercial building under regular operation, which is subject to significant disturbances such

Department of Mechanical Engineering, University of California, Berkeley, USA. datong.zhou@berkeley.edu

Dept. of El. Engineering and Computer Sciences, University of California, Berkeley, USA. [qiehu, tomlin]@eecs.berkeley.edu

This work is supported in part by NSF under CPS:ActionWebs (CNS-0931843) and CPS:FORCES (CNS-1239166).

*These authors contributed equally.

as occupancy. Second, although the existing literature mentions the differences between data-driven and physics-based models, the prevailing isolationist approach does not provide any quantitative comparison with respect to building control applications - a gap we aim to fill by juxtaposing a data-driven with a physics-based model.

The remainder of this paper is organized as follows: In Section II, we describe the testbed and the experimental data. Section III presents the identification process for a purely data-driven model with semiparametric regression, followed by Section IV, which details the procedure for identifying a physics-based model. Section V then compares the performance of both models in terms of open-loop prediction accuracy and closed-loop energy efficient optimal control. We show that, despite the higher accuracy of the complex physics-based model, the optimal control strategies with respect to HVAC operation cost while maintaining the thermal comfort of occupants is almost identical for both systems. Section VI concludes.

II. Preliminaries

A. Testbed for System Identification

We model the temperature evolution of the fourth floor of Sutardja Dai Hall (SDH), a building on the University of California, Berkeley campus. This floor contains offices for research staff and open workspaces for students, and is divided into six zones for modeling purposes (Figure 1).

SDH is equipped with a variable air volume (VAV) HVAC system, which consists of large supply fans driving air through heat exchangers, cooling it down to a desired supply air temperature (SAT), and then distribute air to VAV boxes located throughout the building. There are 21 VAV boxes located on the fourth floor that govern the airflow to each room. In addition, the supply air may be reheated at the VAV box before entering the room.

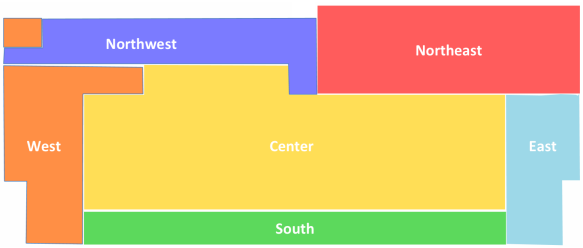

Fig. 1: Zones for the 4th Floor of Sutardja Dai Hall (SDH)

B. Collection of Experimental Data

We collected 51 weeks of one-minute resolution temperature data for the six zones along with the airflow rates of the 21 VAV boxes, SAT, and the outside air temperature. The hourly global horizontal solar radiation data was obtained from a nearby weather station [21], from which the incidence solar radiation of the four geographic directions was calculated with the `PV_LIB` toolbox [22]. All collected data were down-sampled or interpolated to 15 minute intervals.

These 51 weeks of data span periods of normal operation as well as excitation experiments, which were performed in order to increase identifiability of the building model. These experiments were conducted during Saturdays to (a) minimize effects due to occupancy on our collected data, and thus facilitate subsequent parameter identification, (b) to minimize impact on building operation, and (c) to exploit larger comfort bounds on room temperatures during weekend days. Indeed, the comfort bounds were never violated during the forced experiments. Details on the design of our excitation experiments can be found in [18].

C. Data Splitting

Next, we defined the seasons "fall" (early September - mid December), "winter" (mid December - late January), and "spring" (late January - mid May) to account for different occupancy levels during the fall and spring semesters and the winter break. A 90%-10% random split into training and test data was used for fitting and testing the models, respectively.

III. Data-Driven Model

Using semiparametric regression, we identify a difference equation for the temperature evolution, for a lumped zone model and a multi-zone model of the fourth floor of SDH. Semiparametric regression in buildings has been proposed by [15], where the authors used only one week of historic data to model the temperature evolution and used the HVAC's supply air temperature as the single control input. We extend this approach by taking into account multiple weeks, which we separate into three seasons (fall, winter, spring) so as to characterize the different levels of the exogenous heating load for different temporal seasons. In addition, we model the room temperatures as a function of airflow rates from multiple VAVs to obtain a model which can be used for more sophisticated control strategies.

A. Lumped Zone

1) Model Setup: In order to facilitate analysis, the entire 4th floor of SDH is treated as a single zone, with the scalar temperature x corresponding to the average temperature on the entire floor and the input u as the sum of the inflow of all 21 VAV boxes. Then, the temperature evolution is assumed to have the following form:

$$x(k+1) = ax(k) + bu(k) + c^\top v(k) + q_{\text{IG}}(k) + \epsilon(k), \quad (1)$$

where $v := [v_{\text{Ta}}, v_{\text{Ts}}, v_{\text{solE}}, v_{\text{solN}}, v_{\text{solS}}, v_{\text{solW}}]^\top$ is a vector of known disturbances that describes ambient air temperature, the HVAC system's supply air temperature, and solar radiation from each of the four geographic directions. In addition, q_{IG} represents the internal gains due to occupancy and electric devices, and ϵ denotes independent and identically distributed zero mean noise with constant and finite variance which is conditionally independent of x, u, v, and q_{IG}. Finally, a, $b \in \mathbb{R}$ and $c \in \mathbb{R}^6$ are unknown coefficients to be estimated using semiparametric regression [23].

2) Smoothing of Time Series: The q_{IG} term of Equation (1) is treated as a nonparametric term, so that (1) becomes a partially linear model [24]. By taking conditional expectations on both sides of (1), we obtain

$$\hat{x}(k+1) = a\hat{x}(k) + b\hat{u}(k) + c^\top \hat{v}(k) + \mathbb{E}\left[q_{\text{IG}}(k)|k\right] + \mathbb{E}\left[\epsilon(k)|k\right], \quad (2)$$

where the conditional expectations $\hat{x}(\cdot) = \mathbb{E}\left[x(\cdot)|\cdot\right]$, $\hat{u}(\cdot) = \mathbb{E}\left[u(\cdot)|\cdot\right]$, and $\hat{v}(\cdot) = \mathbb{E}\left[v(\cdot)|\cdot\right]$ are used. Noting that $\mathbb{E}\left[\epsilon(\cdot)|\cdot\right] = 0$ and assuming $\mathbb{E}\left[q_{\text{IG}}(\cdot)|\cdot\right] = q_{\text{IG}}(\cdot)$, subtracting (2) from (1) gives

$$x(k+1) - \hat{x}(k+1) = a\left(x(k) - \hat{x}(k)\right) + b\left(u(k) - \hat{u}(k)\right) + c^\top\left(v(k) - \hat{v}(k)\right) + \epsilon(k). \quad (3)$$

The unknown internal gains term has been eliminated, and thus the coefficients a, b, c in (3) can be estimated with any regression method. The conditional expectations $\hat{x}(\cdot), \hat{u}(\cdot)$ and $\hat{v}(\cdot)$ are obtained by smoothing the respective time series [15]. We made use of locally weighted linear regression with a tricube weight function, where we use k-fold cross-validation to determine the optimal kernel width. The error measure used for in-sample estimates is the *Root Mean Squared* (RMS) *Error* between the measured temperatures $\bar{x}(k)$ and the model's predicted temperatures $x(k)$ over a time horizon of N steps (e.g. we chose 24 hours, $N = 96$):

$$\text{RMS error} = \left(\frac{1}{N}\sum_{k=1}^{N}\left[\bar{x}(k) - x(k)\right]^2\right)^{1/2}. \quad (4)$$

3) Bayesian Constrained Least Squares: A major challenge in identifying the model is that commercial buildings are often insufficiently excited. For example, SDH's room temperatures under regular operation only vary within a range of 2°C whereas inflow of the VAV boxes hardly varies at all. To overcome this, forced response experiments (Section II-B) were conducted to compensate for the lack of excitation. Further, we use Bayesian regression, which allows prior knowledge of the building physics to be incorporated in the identification of coefficients. Specifically, Gaussian prior distributions are used for the coefficients a and b, i.e., $a \sim \mathcal{N}(\mu_a, \Sigma_a)$ and $b \sim \mathcal{N}(\mu_b, \Sigma_b)$, where $\mathcal{N}(\mu, \Sigma)$ denotes a jointly Gaussian distribution with mean μ and covariance matrix Σ. In addition, a, b and c are constrained to be identical for the different seasons, since they model the underlying physics of the building which are assumed to be invariant throughout the year.

Let $\mathcal{T} = \{1, 2, \cdots, N\}$ denote N weeks of training data and $\mathcal{F} = \{i \in \mathcal{T}$ such that i is a week in fall$\}$ the set of training weeks from the fall season. Similarly, define the sets of training weeks from the winter and spring as $\mathcal{W}$ and $\mathcal{S}$. The coefficient identification problem reads as follows:

$$\begin{aligned}(\hat{a}, \hat{b}, \hat{c}) = \arg\min_{a,b,c}\ & (J_\mathcal{F} + J_\mathcal{W} + J_\mathcal{S}) + \|\Sigma_a^{-1/2}(a - \mu_a)\|^2 \\ & + \|\Sigma_b^{-1/2}(b - \mu_b)\|^2 \\ \text{s.t. } J_\mathcal{X} = & \textstyle\sum_{i\in\mathcal{X}} \|x_i(k+1) - \hat{x}_i(k+1) - a\left(x_i(k) - \hat{x}_i(k)\right) \\ & - b\left(u_i(k) - \hat{u}_i(k)\right) - c^\top\left(v_i(k) - \hat{v}_i(k)\right)\|^2 \quad (5) \\ & \text{for } \mathcal{X} \in \{\mathcal{F}, \mathcal{W}, \mathcal{S}\}, \\ & 0 < a < 1,\ b \le 0,\ c \ge 0.\end{aligned}$$

$J_\mathcal{F}$, $J_\mathcal{W}$ and $J_\mathcal{S}$ denote the sum of squared errors between actual and predicted temperatures for fall, winter, and spring, respectively. The sign constraints on the parameters b and c capture the fact that temperature correlates positively with all components in v and negatively with VAV airflow. The range of a is a consequence of Newton's Law of Cooling.

To find the effect of the VAV inflow on the 15-minute temperature evolution, we computed the 15-minute incremental reductions in temperature Δx recorded during the excitation experiments. It is assumed that the large inflow u dominates all other effects such that we can assume $\Delta x = x(k+1) - x(k) = b \cdot u(k)$ for all k during the excitation period. The estimated prior μ_b can then be isolated. The prior μ_a was set as the optimal $\hat{a}$ identified by (5) without the prior terms. The covariance matrices Σ_a and Σ_b were chosen subjectively.

4) Estimation of Internal Gains: With the estimated coefficients $\hat{a}, \hat{b}, \hat{c}$ in hand, the internal gains q_{IG} can be estimated by manipulating (2):

$$\hat{q}_{\text{IG}}(k) = \hat{x}(k+1) - \left(\hat{a}\hat{x}(k) + \hat{b}\hat{u}(k) + \hat{c}^\top\hat{v}(k)\right). \quad (6)$$

(6) is used to estimate an instance of the internal gains function for each week i in the training set $\mathcal{T}$. The q_{IG} for each season is defined as the average of estimated weekly gains for all weeks $i \in \mathcal{X}$ and $\mathcal{X} \in \{\mathcal{F}, \mathcal{W}, \mathcal{S}\}$.

5) Results: The estimated internal gains for each season are shown in Figure 2. Observe that, for all three seasons, the

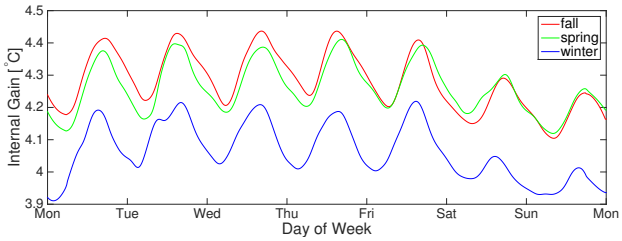

Fig. 2: Estimated Internal Gain q_{IG} from the Data-Driven Model by Season, Lumped Case

internal gains exhibit a daily trend with local peaks around the late afternoon and local minima at night. Moreover, the amplitudes of the internal gains are considerably smaller during weekends, suggesting a lighter occupancy. It can further be seen that the magnitude of the internal gains is smallest for the winter season, which is consistent with intuition as many building occupants are absent.

Lastly, since the Bayesian Constrained Least Squares algorithm (5) has identified a set of parameter estimates $\hat{a}, \hat{b}, \hat{c}$ valid for all three seasons to account for the time-invariant physics of the building, the temperature predictions are of the same nature for all three seasons. We thus conclude that the inherent differences between the seasonal temperature data are captured by the internal gains and can be compared between the seasons on a relative level.

The identified models for the seasons $\mathcal{X} \in \{\mathcal{F}, \mathcal{W}, \mathcal{S}\}$ found with (5) are

$$\begin{aligned}x(k+1) &= 0.80 \cdot x(k) - 0.18 \cdot u(k) \\ &\quad + [0.0019, 0.028, \mathbf{0}]\, v(k) + q_{\text{IG},\mathcal{X}}(k) \\ &= 0.80 \cdot x(k) - 0.18 \cdot u(k) \\ &\quad + 0.0019 \cdot v_{\text{Ta}}(k) + 0.028 \cdot v_{\text{Ts}}(k) + q_{\text{IG},\mathcal{X}}(k)\end{aligned} \quad (7)$$

The estimated effect of solar radiation on the room temperature is orders of magnitude less than that of other factors and hence can be neglected. This is in agreement with our testbed having no windows in the South and Northeast zones, and most of the windows are covered with blinds.

The average RMS prediction errors are 0.22°C, 0.17°C and 0.23°C for fall, winter and spring, respectively, showing that our model predicts the temperature reasonably well.

B. Individual Zones

1) Model Setup: Rather than approximating the entire 4th floor of SDH as a single zone, we now identify a multivariate thermodynamic model for each of the six individual zones:

$$x(k+1) = Ax(k) + Bu(k) + Cv(k) + q_{\text{IG},\mathcal{X}}(k) \quad \text{for } \mathcal{X} \in \{\mathcal{F}, \mathcal{W}, \mathcal{S}\}, \tag{8}$$

where x, $q_{\text{IG},\mathcal{X}} \in \mathbb{R}^6$, and the control input $u \in \mathbb{R}^6$ represent the temperatures, the internal gains of each zone, and the total air flow to each zone, respectively. In the lumped case, it was observed that solar radiation only had a negligible effect on the building's thermodynamics compared to the input and other disturbances, and thus we omit the solar radiation in the subsequent analysis: $v := [v_{\text{Ta}}, v_{\text{Ts}}]^\top \in \mathbb{R}^2$.

Inspired by Newton's Law of Cooling, only adjacent zones influence each other's temperature, which defines the sparsity pattern of the coefficient matrices that are to be estimated:

$$A_{ij} = \begin{cases} \neq 0, & \text{if } i = j \text{ or } (i,j) \text{ adjacent} \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

The diagonal elements of A denote autoregressive terms for zone temperatures, whereas non-diagonal elements describe the heat exchange between adjacent rooms. The matrix B is diagonal by definition of u. The sparsity pattern of C is found by physical adjacency of a respective zone to an exterior wall of a given geographic direction.

2) Model Identification: The procedure for the estimation of the parameter matrices $\hat{A}$, $\hat{B}$, $\hat{C}$, and the internal gains follows (5), but with a modified choice of the (now matrix-valued) priors μ_a and μ_b: μ_b and the diagonal entries of μ_a are obtained by scaling the corresponding priors from the lumped zone case in order to account for the thermal masses of the individual zones, which are smaller than in the lumped case. The off-diagonal elements of μ_a, which represent the heat transfer between adjacent zones, were set to a value close to zero, according to our calculations with the heat transfer equation $\dot{q} = U \cdot A \cdot \Delta x$ and [25].

3) Results: Figure 3 shows the estimated internal gains for the three seasons fall, winter, and spring for the six single zones, computed with the smoothed time series (6). It can be seen that the different zones exhibit different magnitudes of internal gains, with average values of the internal gains ranging between 1.0°C and 3.6°C for different zones and seasons. Similar to the lumped zone case (Figure 2), daily peaks of the internal gains profiles can be recognized, with a slight decrease in magnitude on weekend days. Table I reports the average prediction RMS error by zone and season.

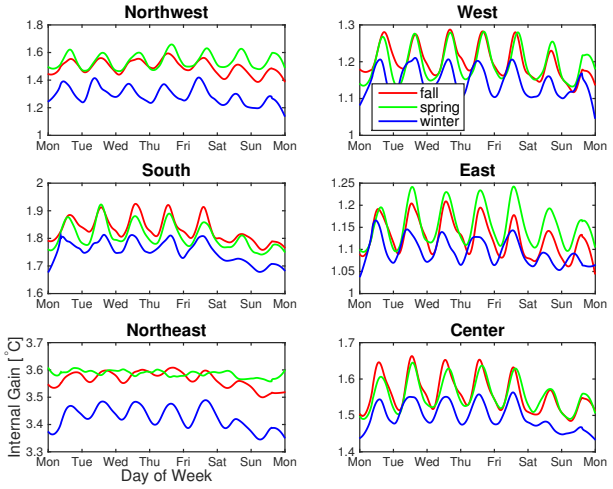

Fig. 3: Estimated Internal Gain q_{IG} from the Data-Driven Model by Zone and Season, Individual Case

Data-Driven Model							
Season	NW	W	S	E	NE	C	Mean
Fall	0.98	0.61	0.28	0.42	0.28	0.36	0.488
Winter	1.41	0.34	0.29	0.26	0.25	0.21	0.460
Spring	0.56	0.25	0.31	0.71	0.17	0.34	0.390

Physics-Based Model							
Season	NW	W	S	E	NE	C	Mean
Fall	0.61	0.46	0.39	0.39	0.20	0.32	0.396
Winter	0.55	0.39	0.34	0.32	0.18	0.24	0.338
Spring	0.45	0.28	0.24	0.33	0.09	0.19	0.263

TABLE I: RMS by Zone and Season for Data-Driven and Physics-Based Models

IV. Physics-Based Model

We describe the physics-based modeling approach proposed in [18], which obtains an RC-model using the Building Resistance-Capacitance Modeling (BRCM) MATLAB toolbox [26]. We re-identify the building model using the same training dataset as used in Section III, and estimate distinct internal gains functions for different seasons.

A. Model Setup

The physics-based model has the following form [18]:

$$x(k+1) = Ax(k) + B_v v(k) + B_{\text{IG}} f_{\text{IG}}(k) + \textstyle\sum_{i=1}^{21} \left(B_{xu_i} x(k) + B_{vu_i} v(k)\right) u_i(k) \tag{10a}$$

$$y(k) = Cx(k), \tag{10b}$$

where state vector $x \in \mathbb{R}^{289}$ and $y \in \mathbb{R}^6$ represent temperatures of all building elements (walls, ceilings, floors, etc.) on the 4th floor and the average temperatures of the six zones shown in Figure 1, respectively. $u \in \mathbb{R}^{21}$ denotes the airflow rate from the 21 VAV boxes and $v := [v_{\text{Ta}}, v_{\text{Ts}}]^\top$ the vector of known disturbances. As in the data-driven model, heat gains due to solar radiation are omitted from the analysis. Finally, $f_{\text{IG}}(k) : \mathbb{N} \to \mathbb{R}^6$ captures internal gains in each of the six zones on the 4th floor. For week m in the training set $\mathcal{T}$:

$$f_{\text{IG}}(k) = f^c_{\text{IG}} + \begin{cases} f^v_{\text{IG},\mathcal{F}}(k), & \text{if } m \in \mathcal{F}, \\ f^v_{\text{IG},\mathcal{W}}(k), & \text{if } m \in \mathcal{W}, \\ f^v_{\text{IG},\mathcal{S}}(k), & \text{if } m \in \mathcal{S}, \end{cases} \tag{11}$$

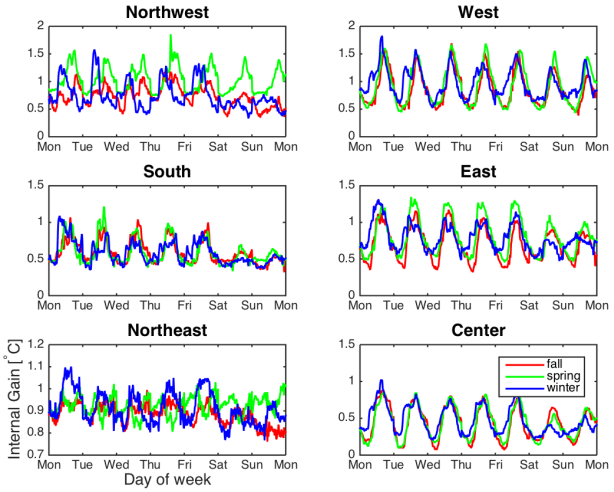

Fig. 4: Estimated Internal Gain f_{IG} from the Physics-Based Model by Zone and Season

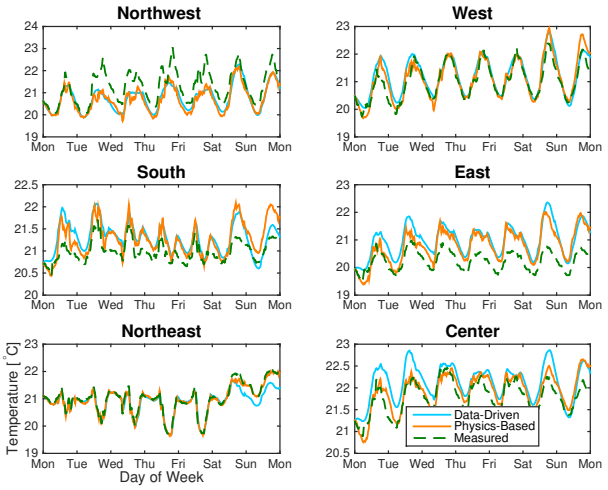

Fig. 5: Simulated Temperatures from the Data-Driven Model (blue), Physics-Based Model (orange) and Actual Temperatures (green)

where f_{IG}^c is an unknown constant vector representing background heat gains due to idle electric appliances. Functions $f_{\text{IG},\mathcal{F}}^v(\cdot)$, $f_{\text{IG},\mathcal{W}}^v(\cdot)$ and $f_{\text{IG},\mathcal{S}}^v(\cdot)$ are unknown nonparametric functions that capture the time-varying heat gain due to occupancy and equipments in fall, winter and spring, respectively. The system matrices A, B_v, B_{IG}, B_{xu_i} and B_{vu_i} are functions of the window heat transmission coefficient U_{win} and convection coefficients of the interior wall γ_{IW}, exterior wall γ_{EW}, floor γ_{floor}, and ceiling γ_{ceil}. Define $\gamma := \left[U_{\text{win}}, \gamma_{\text{IW}}, \gamma_{\text{EW}}, \gamma_{\text{floor}}, \gamma_{\text{ceil}}, f_{\text{IG}}^{c\top}\right]^\top \in \mathbb{R}^{11}$, then to identify the physics-based model, we need to estimate the parameter vector γ as well as the functions $f_{\text{IG},\mathcal{X}}^v(\cdot)$, $\mathcal{X} \in \{\mathcal{F}, \mathcal{W}, \mathcal{S}\}$.

B. Model Identification

For a fair comparison, the same data used to train and test the data-driven model is used to train and validate the physics-based model. The model identification process is performed in two steps: First, the subset of the training data collected during weekends is used to estimate the parameters, γ. Second, the nonparametric functions $f_{\text{IG},\mathcal{X}}^v(\cdot)$ are estimated from the complete training dataset.

1) Parameter Estimation: We first set $f_{\text{IG},\mathcal{X}}^v(\cdot) = 0$ during the weekend days to evaluate them at a later point. With $f_{\text{IG},\mathcal{X}}^v(\cdot) = 0$, (10) reduces to a purely parametric model:

$$\begin{aligned} x(k+1) &= Ax(k) + B_v v(k) + B_{\text{IG}} f_{\text{IG}}^c \\ &\quad + \textstyle\sum_{i=1}^{21} \left(B_{xu_i} x(k) + B_{vu_i} v(k)\right) u_i(k), \\ y(k) &= Cx(k). \end{aligned} \tag{12}$$

The optimal model parameters are estimated by solving the following optimization problem:

$$\begin{aligned} \hat{\gamma} = \ & \arg\min_{\gamma > 0} \ \textstyle\sum_{m \in \mathcal{T}} \sum_k \|y_m(k,\gamma) - \bar{y}_m(k)\|^2 \\ \text{s.t.} \ & \ y_m(k,\gamma) \text{ and } x_m(k,\gamma) \text{ satisfy (12) with} \\ & \ x_m(0) = x_{\text{KF},m}(0) \\ & \ u_m(k) = \bar{u}_m(k), v_m(k) = \bar{v}_m(k) \ \forall \ k, \end{aligned} \tag{13}$$

where $\bar{u}$, $\bar{v}$ and $\bar{y}$ denote the measured inputs, disturbances, and zone temperatures, respectively. In other words, we choose γ such that, when the model is simulated with this set of parameter values and the measured inputs and disturbances, the sum of squared errors between the measured zone temperatures and the simulated temperatures is minimized. The initial state $x_m(0)$ is required to simulate the model, however, not all states are measurable (e.g. wall temperature). Thus, we estimate the initial states using a Kalman Filter $x_{\text{KF},m}(0)$ and set $x_m(0) = x_{\text{KF},m}(0)$. Furthermore, to compensate for the lack of sufficient excitation of the building, physically plausible initial guesses for γ are chosen. The optimal parameter values are similar to those reported in [18] and hence omitted due to space limitations.

2) Estimation of $f_{IG}^v(\cdot)$: Let $f_{\text{IG},m}^v(\cdot)$ be an instance of the internal gains function $f_{\text{IG}}^v(\cdot)$ estimated for week m in the training set. The optimal estimate for a given season, is then defined as the the average of all estimates for that season:

$$\hat{f}_{\text{IG},\mathcal{F}}^v(k) = \textstyle\sum_{m \in \mathcal{F}} f_{\text{IG},m}^v(k) / \|\mathcal{F}\| \quad \forall \ k, \tag{14}$$

where $\|\mathcal{F}\|$ represents the cardinality of set $\mathcal{F}$. To estimate $f_{\text{IG},m}^v(\cdot)$ for a given week m, let $\tilde{x}(k)$ and $\tilde{y}(k)$ denote the predicted states and zone temperatures at time k, with $f_{\text{IG},w}^v(k-1) = 0$. That is,

$$\begin{aligned} \tilde{x}(k) &= Ax(k-1) + B_v v(k-1) + B_{\text{IG}} f_{\text{IG}}^c \\ &\quad + \textstyle\sum_{i=1}^{21} \left(B_{xu_i} x(k-1) + B_{vu_i} v(k-1)\right) \\ &\quad \cdot u_i(k-1), \\ \tilde{y} &= C\tilde{x}(k). \end{aligned} \tag{15}$$

By noting $x(k) = \tilde{x}(k) + B_{\text{IG}} f_{\text{IG},m}^v(k-1)$, $f_{\text{IG},m}^v(k-1)$ can be estimated by solving $(CB_{\text{IG}}) \cdot f_{\text{IG},m}^v(k-1) = \bar{y}(k) - \tilde{y}(k)$, a set of linear equations, using Ordinary Least Squares. $\bar{y}(k)$ denotes the measured zone temperatures at time k.

C. Results

The average daily prediction RMS errors by zone and season are reported in Table I. Figure 4 shows the estimated temperature contribution of the internal gains for fall, winter and spring. The zones that correspond to open workspaces and conference rooms ("West", "South", "East", "Center")

show discernible daily peaks in their internal gains profiles with a slight decrease during weekends. Lastly, there is little variation in the internal gains across different seasons.

V. Quantitative Comparison of both Models

A. Prediction Accuracy

The physics-based model (Model B) is found to have a higher prediction accuracy compared to the data-driven model for the individual zones (Model A) presented in Section III-B: According to Table I, the mean RMS error for Model B across zones is $0.11°\text{C}$ lower than for Model A. This is also illustrated in Figure 5, which shows 7-day open-loop predictions of the temperature of a randomly selected holdout test week in the spring period, simulated with both models initialized with the measured temperature. To the best of our knowledge, a quantitative comparison at this level is non-existent, as previous building models were developed for different testbeds, fictitious buildings or from simulated data. This paper attempts to close this gap by providing a quantitative comparison between the low-dimensional data-driven model and the high-dimensional physics-based model's prediction accuracy for the *same* multi-zone commercial building under *regular* operation.

B. Energy Efficient Control

Next, we explore the extent to which Model A's slightly lower prediction accuracy affects its resulting controller's closed-loop performance in energy efficient control. We formulate an MPC problem to find the optimal control strategy that minimizes the cost of HVAC operation over the same week used in Figure 5, while guaranteeing the temperature to stay within a comfort zone $[T_{\min}, T_{\max}]$, which we chose as $[20°\text{C}, 22°\text{C}]$ [27], and confining the control input to the minimum and maximum airflow settings of the HVAC system $[u_{\min}, u_{\max}]$:

$$\begin{aligned}
\min_{u,\varepsilon} \quad & \sum_{k=1}^{N} u(k)^2 + \rho\|\varepsilon\|_2 \\
\text{s.t. } & x(0) = \bar{x}(0) \\
& x(k+1) = \begin{cases} (8), & \text{Model A} \\ (10a), & \text{Model B} \end{cases} \qquad (16) \\
& u_{\min} - \varepsilon \le u(k) \le u_{\max} + \varepsilon \qquad \forall k \in [0, N-1] \\
& \begin{cases} T_{\min} \le x(k) \le T_{\max}, & \text{Model A} \\ T_{\min} \le Cx(k) \le T_{\max}, & \text{Model B} \quad (10b) \end{cases} \quad \forall k \in [1, N]
\end{aligned}$$

The temperature is initialized with the measured temperature $\bar{x}(0)$ at the beginning of the week-long simulation. Soft constraints on the control input with a penalty parameter ρ ensure feasibility of the problem. The penalty represents the cost of decreasing airflow below the set minimum value. This is physically feasible as the set minimum airflow rate for our testbed is significantly higher than the standard minimum required by building standards. To find the optimal control strategy, we make use of receding horizon control with a prediction horizon of three 15-minute time steps.

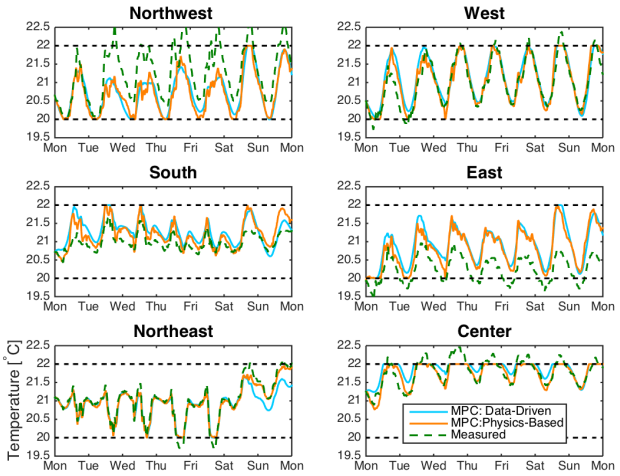

Fig. 6: Optimal Temperature for MPC with Data-Driven Model (blue), MPC with Physics-Based Model (orange) and Actual Temperature (green)

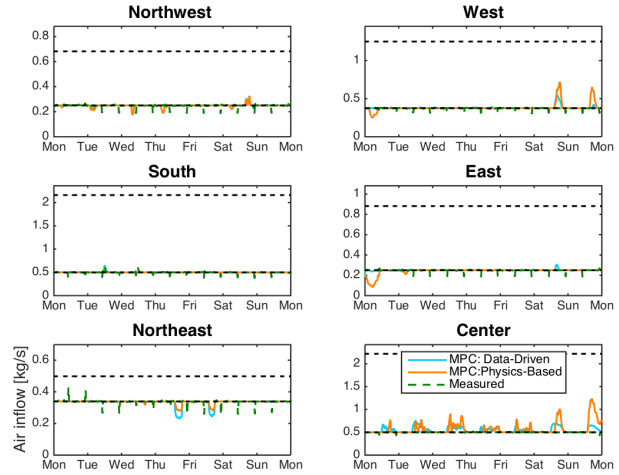

Fig. 7: Optimal Control Strategy for MPC with Data-Driven Model (blue), MPC with Physics-Based Model (orange) and Actual Input (green)

Figure 6 shows the temperature trajectory computed by the energy efficient controller (16) computed with both models A and B, together with the measured temperature as a reference. It can be seen that both control schemes are capable of maintaining the temperature within $[20°\text{C}, 22°\text{C}]$, with a control strategy that is of comparable cost (1,006 and 1,731 for Model A and Model B, respectively, where $\rho = 100$), shown in Figure 7. Furthermore, it is interesting to observe that variations in the control input do not impact the periodicity of the temperature qualitatively, which can be explained by the regularity of the identified internal gains.

These findings suggest that both models perform equally well in designing an energy efficient control strategy. However, computing this strategy for Model A was cheap (< 5 minutes) compared to Model B (≈ 20 hours) on a 2 GHz Intel Core i7, 16 GB 1600 MHz DDR3 machine. Further, we note that in real-world applications, the MPC would use state feedback to initialize the temperature with sensor measurements at every time step, whereas in our simulation, it operates in an "open loop" fashion and hence propagates the estimation error with time. This will reduce the difference in the prediction quality by both controllers even further,

since the RMS error is now to be evaluated on a much shorter prediction horizon, thereby further corroborating the finding of almost identical control schemes.

Observing that Model A only suffers a negligible loss of accuracy compared to Model B for an open loop optimal control scheme, our findings suggest the suitability of Model A to applications with temperature-critical zones in which even more precise temperature estimates are needed, e.g. long-term planning of reserve provision for frequency regulation.

VI. Discussion and Conclusion

We identified a low-dimensional data-driven model, using semiparametric regression, and a high-dimensional physics-based resistance-capacitance model for the thermal behavior of the same multi-zone commercial building. Both state-space models were fitted on experimental data collected during regular building operation and capture the effect of disturbances such as occupancy and electrical appliances that commercial buildings are subjected to, without installation of any additional hardware such as occupancy sensors.

The identification of both models on the *same building* enabled us to quantitatively compare the performance of these types of models when applied to a real building, which has not been investigated before. Our results showed that the RMS error of the open-loop temperature prediction of the physics-based model across different thermal zones and temporal seasons is $0.11°\text{C}$ lower than in the data-driven model, a 25% reduction. However, simulating energy efficient MPC schemes under both models suggested both models perform equally well in terms of cost function minimization and constraint satisfaction despite the significantly higher complexity of the physics-based model.

It is widely known in this field that low-dimensional data-driven models have lower prediction accuracy than high-dimensional physics-based models, and thus have been only proposed for control of less temperature-critical buildings or zones. However, our work investigated an identification method for data-driven models for multi-zone commercial buildings in regular operation and demonstrated that the lower open-loop prediction accuracy of such data-driven models is insignificant in closed-loop control schemes compared to a high-dimensional physics-based model. Based on these findings, we suggest that such data-driven models may be suitable for applications that were previously considered inappropriate, e.g. frequency regulation.

Finally, we are currently working on verifying our hypothesis by designing and implementing a control scheme suitable for frequency regulation, using the data-driven model, into the building operation system of SDH.

Acknowledgment

We thank Anil Aswani for fruitful discussions.

References

[1] L. Pérez-Lombard, J. Ortiz, and C. Pout, "A Review on Buildings Energy Consumption Information," *Energy and Buildings*, no. 40, pp. 394–398, 2008.

[2] U.S. Department of Energy Buildings Energy Data Book. [Online]. Available: http://buildingsdatabook.eren.doe.gov/

[3] J. Široky, F. Oldewurtel, J. Cigler, and S. Prívera, "Experimental Analysis of Model Predictive Control for an Energy Efficient Building Heating System," *Applied Energy*, vol. 88, pp. 3079–3087, 2011.

[4] A. Parisio, L. Fabietti, M. Molinari, D. Varagnolo, and K. H. Johansson, "Control of HVAC Systems via Scenario-Based Explicit MPC," *IEEE Conference on Decision and Control*, 2014.

[5] M. Balandat, F. Oldewurtel, M. Chen, and C. Tomlin, "Contract Design for Frequency Regulation by Aggregations of Commercial Buildings," *52nd Annual Allerton Conference on Communication, Control, and Computing*, 2014.

[6] E. Vrettos, F. Oldewurtel, F. Zhu, and G. Andersson, "Robust Provision of Frequency Reserves by Office Building Aggregations," *Proceedings of the 19th IFAC World Congress*, pp. 12 068 – 12 073, 2014.

[7] Y. Lin, P. Barooah, S. Meyn, and T. Middelkoop, "Experimental Evaluation of Frequency Regulation From Commercial Building HVAC Systems," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, 2015.

[8] F. Baccino, F. Conte, S. Massucco, F. Silvestro, and S. Grillo, "Frequency Regulation by Management of Building Cooling Systems through Model Predictive Control," *Power Systems Computation Conference*, pp. 1–7, 2014.

[9] M. Maasoumy, M. Razmara, M. Shahbakhti, and A. Sangiovanni-Vincentelli, "Handling Model Uncertainty in Model Predictive Control for Energy Efficient Buildings," *Energy and Buildings*, 2014.

[10] H. Hao, A. Kowli, Y. Lin, P. Barroah, and S. Meyn, "Ancillary Service for the Grid via Control of Commercial Building HVAC Systems," *American Control Conference*, no. 467-472, June 2013.

[11] M. Duffy, M. Hiller, D. Bradley, W. Keilholz, and J. Thornton, "TRNSYS - Features and Functionality for Building Simulation," *IBSPA Conference*, pp. 1950 – 1954, 2009.

[12] J. Zhao, K. P. Lam, and B. E. Ydstie, "EnergyPlus model-based predictive control (EPMPC) by using MATLAB/SIMULINK and MLE+," *Proceedings of 13th Conference of International Building Performance Simulation Association*, 2013.

[13] J. R. Dobbs, "A Comparison of Thermal Zone Aggregation Methods," *51st Conference on Decision and Control*, 2012.

[14] S. Goyal and P. Barooah, "A Method for Model-Reduction of Nonlinear Thermal Dynamics of Multi-Zone Buildings," *Energy and Buildings*, no. 47, pp. 332–340, 2012.

[15] A. Aswani, N. Master, J. Taneja, V. Smith, A. Krioukov, D. Culler, and C. Tomlin, "Identifying Models of HVAC Systems Using Semiparametric Regression," *American Control Conference*, 2012.

[16] P. Radecki and B. Hencey, "Online Building Thermal Parameter Estimation via Unscented Kalman Filtering," *American Control Conference*, 2012.

[17] ——, "Online Thermal Estimation, Control, and Self-Excitation of Buildings," *52nd Conference on Decision and Control*, 2013.

[18] Q. Hu, F. Oldewurtel, M. Balandat, E. Vrettos, D. Zhou, and C. Tomlin, "Model Identification of Commercial Building HVAC Systems During Regular Operation - Empirical Results and Challenges," *American Control Conference (Accepted)*, 2016.

[19] W. J. Cole, E. T. Hale, and T. F. Edgar, "Building Energy Model Reduction for Model Predictive Control Using OpenStudio," *American Control Conference*, 2013.

[20] S. Goyal, H. A. Ingley, and P. Barooah, "Occupancy-Based Zone-Climate Control for Energy-Efficient Buildings: Complexity vs. Performance," *Applied Energy*, no. 106, pp. 209–221, 2013.

[21] "CIMIS Station Reports," California Irrigation Management Information System, Tech. Rep., 2015. [Online]. Available: http://www.cimis.water.ca.gov/

[22] "PV Performance Modeling Collaborative." [Online]. Available: https://pvpmc.sandia.gov/

[23] D. Ruppert, M. Wand, and R. Carroll, *Semiparametric Regression*. Cambridge University Press, 2003.

[24] W. Härdle, H. Liang, and J. Gao, *Partially Linear Models*. Springer, 2000.

[25] S. Koehler and F. Borrelli, "Building Temperature Distributed Control via Explicit MPC and "Trim and Respond" Methods," *European Control Conference*, 2013.

[26] D. Sturzenegger, D. Gyalistras, M. Morari, and R. Smith, "Semi-Automated Modular Modeling of Buildings for Model Predictive Control," *BuildSys 2012 – Workshop of SCM SenSys Conference*, 2012.

[27] S. J. Hansen and H. Burroughs, *Managing Indoor Air Quality*. Lulu Press, Inc., 2013.