Model Comparison of a Data-Driven and a Physical Model for Simulating HVAC Systems

Datong 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 the provision of frequency regulation to support grid stability. To realize energy savings and frequency regulation, it is of prime importance to identify a realistic model for the temperature dynamics of a building. We identify a low-dimensional data-driven model and a high-dimensional physics-based model for different spatial granularities and temporal seasons based on a case study of an entire floor of Sutardja Dai Hall, an office building on the University of California, Berkeley campus. A comparison of these contrasting models shows that, despite the higher forecasting accuracy of the physics-based model, both models perform almost equally well for energy efficient control. We conclude that the data-driven model is more amenable to controller design due to its low complexity, and could serve as a substitution for highly complex physics-based models with an insignificant loss of prediction accuracy for many applications. On the other hand, our physics-based approach is more suitable for modeling buildings with finer spatial granularities.

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]. More recently, it has been proposed to engage buildings in supporting the supply quality of electricity and the grid stability, by participating in the regulation of electricity’s frequency [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], [11], [12], TRNSYS [13] and EnergyPlus [14]. These models are motivated by the thermodynamics of the building and explicitly model the heat transfer between components of the buildings. The advantage of such models is their high granularity of temperature modeling, but a drawback is their high dimensionality which makes them computationally expensive. On the other hand, a new direction of research attempts to identify lower-dimensional, data-driven models, e.g. with Input-Output models [6] and semiparametric regression [15]. The purpose is to alleviate the computational complexity in expense for coarser and less accurate temperature predictions.

A crucial question that arises within these two extremes is the extent to which the estimated temperature model is compatible with controller design. Take Model Predictive Control (MPC) for example, where the classical physics-based models require an MPC strategy to be solved online with high computational demand. Even then, the inherent bilinearity ensuing from the physics of the HVAC system often requires robustification, in the form of stochastic MPC formulations with chance constraints [16]. In contrast, regression-based models provide convenient difference equations that are easy and fast to use for MPC. A logical question to ask is how lean a model can be for a reasonable control application, without trading off too much accuracy and granularity of the temperature predictions.

To the best of our knowledge, data-driven and physical models for the identification of temperature evolution in commercial buildings have only been studied in isolation and on individual testbeds (e.g. [17], [3], [6], [18]). The identified models are often validated using simulation data [11] or experimental data collected under controlled environments, e.g. without occupants [6]. This makes a comparison of the performance between these two types of models, for a real building under normal operation, impossible due to their different nature.

In this paper, we identify a data-driven model, using semiparametric regression, and adapt our previous physics-based model for the same building [18], using a one-year period of experimental data. We provide a quantitative comparison using various metrics, including open-loop prediction accuracy and closed-loop control strategies. We show that, despite the higher accuracy of the complex physics-based model compared to the low-dimensional data-driven model, the optimal control strategy with respect to HVAC operation cost while maintaining the thermal comfort of occupants is almost identical for both systems. This indicates that the data-driven model provides enough accuracy for controller design, unless controllers that must maintain building temperatures more accurately or with finer spatial granularities are needed. Finally, a qualitative analysis of the advantages and disadvantages of each type of model, together with their suitability for different applications is provided.

The remainder of this paper is organized as follows: In
Section II, we describe the testbed and the data collected for our research. 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 will then compare the performance of the data-driven model and the physical model under different metrics. We conclude in Section VI with a summary of our current and intended future work.

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).

The building is equipped with a variable air volume (VAV) HVAC system that is common to 30% of all U.S. commercial buildings [19]. The system contains large supply fans which drive 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)](image)

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 from the simple Measurement and Actuation Profile (sMAP). sMAP is a protocol that collects, stores and publishes time-series data from a wide variety of sensors [20], [21]. The hourly global horizontal solar radiation data recorded at a nearby weather station is obtained [22], from which the incidence solar radiation of the four geographic directions is calculated with the PVLIB toolbox [23]. All collected data are down-sampled or interpolated, respectively, to 15 minute intervals.

These 51 weeks of data span periods when the building was under normal operation as well as periods with excitation experiments. For accurate parameter identification, temperatures of neighboring zones should not have strong correlation [24]. Our testbed is a regular office building in operation, thus forced response experiments were performed during Saturdays to (a) increase identifiability of the building model; (b) minimize effects due to occupancy on our data, and thus facilitate subsequent parameter identification; (c) minimize disturbance to building operation [18].

C. Data Splitting

Next, we define the seasons “fall” (early September until mid December), “winter” (mid December until late January), and “spring” (late January until mid May) in order to account for different occupancy levels during the fall and spring semesters, and the winter break. After the weeks have been assigned to the seasons, a random portion of the data in each season (e.g. we chose 90%) is defined as the training data, and the remaining weeks to be removed prior to the analysis are declared as the test set, which will be used to assess the accuracy of the optimal temperature model fitted on the training data.

III. DATA-DRIVEN MODEL

We identify a difference equation for the temperature evolution with semiparametric regression, using the collected data from the 4th floor of SDH. Semiparametric regression in buildings has been proposed by [15], where the authors chose one week of data to model the temperature evolution including an exogenous heating load that captures the effect of occupancy, electric devices, VAV supply air temperature, outside air temperature, and solar radiation. 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.

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 area-averaged zone temperatures and the input $u$ as the sum of the inflow of all 21 VAV boxes. This lumped model assumes a uniform temperature on the entire floor, $x$, and has been commonly used in literature [16], [25]. Then, the temperature evolution is assumed to have the following form:

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

where $u$ denotes the total air inflow to the entire floor and $v := [v_{\text{Ta}}, v_{\text{Tb}}, v_{\text{solE}}, v_{\text{solN}}, v_{\text{solS}}, v_{\text{solW}}]^\top$ the vector of known disturbances that describe ambient air temperature, the HVAC system’s supply air temperature and solar radiation from each of the four geographical directions. In addition, $q_{\text{IG}}$ represents the internal gains due to occupancy and electric devices, and $\epsilon$ denotes independent and identically distributed zero mean noise with constant and finite variance which is conditionally independent of $x$, $u$, $v$, and $q_{\text{IG}}$.

2) Smoothing of Time Series: The unknown coefficients $a, b$, and $c$ are to be estimated using semiparametric regression [26], [27]. The $q_{\text{IG}}$ term of Equation (1) is treated as a nonparametric term, so that (1) becomes a partially linear model. By taking conditional expectations on both sides of (1), we obtain...
\[ \hat{x}(k+1) = a \hat{x}(k) + b \hat{u}(k) + c^T \hat{v}(k) + \mathbb{E}[q_{IG}(k)|k] + \mathbb{E}[e(k)|k], \]

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

\[ x(k+1) - \hat{x}(k+1) = a(x(k) - \hat{x}(k)) + b(u(k) - \hat{u}(k)) + c^T(v(k) - \hat{v}(k)) + \epsilon(k). \]

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 \( \hat{x}(\cdot), \hat{u}(\cdot), \hat{v}(\cdot) \) 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 bandwidth for regression. The error measure used for in-sample estimates is the Root Mean Squared (RMS) Error between the measured temperatures \( \hat{x}(k) \) and the model’s predicted temperatures \( x(k) \) over a time horizon of \( N \) steps (e.g. we chose a 24 hour time horizon, \( N = 96 \)):

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

3) Bayesian Constrained Least Squares: A main challenge in identifying the model is that commercial buildings are often insufficiently excited. Take SDH for example, whose room temperatures under regular operation only vary within a range of 2°C and inflow of the single VAV boxes hardly vary at all. To overcome this, data collected during forced response experiments described in Section II-B was used in training the model. To further compensate for the lack of excitation, a Bayesian regression method is used, which allows our prior knowledge of the building physics to be incorporated in the identification of coefficients. More 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 \( \mu \) and covariance matrix \( \Sigma \). 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. Therefore the coefficient identification problem is formulated as follows:

\[
\begin{align*}
(\hat{a}, \hat{b}, \hat{c}) &= \arg \min_{a, b, c} (J_1 + J_2 + J_3) + \|\Sigma_a^{-1/2}(a - \mu_a)\|^2 \\
&\quad + \|\Sigma_b^{-1/2}(b - \mu_b)\|^2 \\
&\quad \text{s.t. } J_1 = \|x_i(k+1) - \hat{x}_i(k+1) - a(x_i(k) - \hat{x}_i(k)) \\
&\quad \quad - b(u_i(k) - \hat{u}_i(k)) - c^T(v_i(k) - \hat{v}_i(k))\|^2 \\
&\quad \quad \text{for } i \in \{f, w, s\}, \\
&\quad \quad 0 < a < 1, \quad b \leq 0, \quad c \geq 0,
\end{align*}
\]

where subscripts \( f, w, \) and \( s \) represent fall, winter and spring, respectively. The sign constraints on the parameters \( b \) and \( c \) translate into the fact that the temperature to be estimated positively correlates with all components in \( v \) and negatively correlates with the 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 decreases in temperature \( \Delta x \) 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 \( \mu_b \) can then be isolated from (6). The prior \( \mu_a \) was set as the optimal \( \hat{a} \) identified by (5) without the prior terms. The covariance matrices \( \Sigma_a \) and \( \Sigma_b \) 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):

\[ \hat{q}_{IG}(k) = \hat{x}(k+1) - \left( \hat{a} \hat{x}(k) + \hat{b} \hat{u}(k) + \hat{c}^T \hat{v}(k) \right). \]

This can be interpreted as the difference between the smoothed temperature \( \hat{x}(k+1) \) and the predicted expected temperature \( \mathbb{E}[x(k+1)] \). With the estimated parameter coefficients being constant over the different seasons, a distinct function of internal gains is estimated for each season by averaging the estimated weekly gains for a given season.

5) Results: The estimated internal gains for each season, calculated with (7), are shown in Figure 2. Observe that, for all three seasons, the 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 the 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 in accordance with our intuition since most building occupants are absent during that period.

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 different seasons found with (5) are
\[
\begin{align*}
x(k + 1) &= 0.80 \cdot x(k) - 0.18 \cdot u(k) + [0.0019, 0.028, 0] v(k) + q_{IG,i}(k) \\
&= 0.80 \cdot x(k) - 0.18 \cdot u(k) + 0.0019 \cdot v_{Te}(k) + 0.028 \cdot v_{Tc}(k) + q_{IG,i}(k)
\end{align*}
\]
for \( i \in \{f, w, s\} \) \hspace{1cm} (8)

The estimated coefficients of \( c \) corresponding to the solar radiation disturbances are very small \((< 10^{-6})\) compared to the other estimated coefficients. Since the temperatures are of the order 10°C, air inflow around 1 kg/s and solar radiation about 100 W/m², the effect of solar radiation on the room temperature is orders of magnitude less than that of other factors and hence can be neglected.

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, in this section, we identify a multivariate model that describes the thermodynamic behavior of each of the six individual zones:

\[
x(k + 1) = Ax(k) + Bu(k) + Cv(k) + q_{IG,i}(k)
\]

for \( i \in \{f, w, s\} \) \hspace{1cm} (9)

where \( x, q_{IG,i} \in \mathbb{R}^6 \), 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: \( v := [v_{Te}, v_{Tc}]' \in \mathbb{R}^2 \).

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. Hence

\[
A_{ij} = \begin{cases} 
1 & \text{if } i = j \text{ or } (i, j) \text{ adjacent} \\
0 & \text{otherwise.}
\end{cases}
\]

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 \( A, B, C \), and the internal gains follows (5), but with a modified choice of the (now matrix-valued) priors \( \mu_A \) and \( \mu_B \); \( \mu_B \) and the diagonal entries of \( \mu_A \) are obtained by scaling the corresponding priors from the lumped zone case in order to account for the thermal mass of the individual zones, which is smaller than in the lumped case. The off-diagonal elements of \( \mu_A \), 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 \( q = U \cdot A \cdot \Delta x \) and [28].

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 (7). 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. The average prediction RMS error by zone and season are reported in Table I.

![Figure 3. Estimated Internal Gain \( q_{IG} \) from the Data-Driven Model by Zone and Season, Individual Case](image)

| Zone | Fall | Winter | Spring |
|------|------|--------|--------|
| NW   | 0.58 | 0.61   | 0.56   |
| W    | 0.56 | 0.62   | 0.50   |
| S    | 0.61 | 0.52   | 0.61   |
| E    | 0.58 | 0.58   | 0.57   |
| NE   | 0.58 | 0.52   | 0.56   |
| C    | 0.56 | 0.61   | 0.61   |

### Table I

| Zone | Fall | Winter | Spring |
|------|------|--------|--------|
| NW   | 0.39 | 0.46   | 0.39   |
| W    | 0.38 | 0.42   | 0.34   |
| S    | 0.37 | 0.34   | 0.37   |
| E    | 0.36 | 0.34   | 0.36   |
| NE   | 0.36 | 0.34   | 0.36   |
| C    | 0.36 | 0.34   | 0.36   |

### Table II

| Zone | Fall | Winter | Spring |
|------|------|--------|--------|
| NW   | 0.30 | 0.33   | 0.30   |
| W    | 0.30 | 0.33   | 0.30   |
| S    | 0.30 | 0.33   | 0.30   |
| E    | 0.30 | 0.33   | 0.30   |
| NE   | 0.30 | 0.33   | 0.30   |
| C    | 0.30 | 0.33   | 0.30   |

IV. PHYSICS-BASED MODEL

In this section, we describe the physics-based modeling approach proposed in [18], which is a Resistance-Capacitance (RC) model obtained using the Building Resistance-Capacitance Modeling (BRCM) MATLAB toolbox [11]. A main advantage of this approach is that the resulting model has a small number of parameters, even for a
complex multi-zone building; furthermore, these parameters have strong physical meaning, which aids in their identification.

In this paper, 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 building model has the following form [18]:

\[
\begin{align*}
    x(k+1) &= Ax(k) + B_u v(k) + B_{IG} f_{IG}(k) + \sum_{i=1}^{21} (B_{ux}, x(k) + B_{vu}, v(k)) u_i(k) \\
    y(k) &= C x(k),
\end{align*}
\]

where the state vector \( x \in \mathbb{R}^{289} \) represents temperatures of all building elements on the 4th floor and \( y \in \mathbb{R}^6 \) represents the average temperatures of the six zones shown in Figure 1. \( u \in \mathbb{R}^{21} \) denotes the air inflow rates, whose \( i \)-th element \( u_i \) denotes the inflow rate from the \( i \)-th VAV box. \( v := [v_T s, v_T s]^T \) is the disturbance vector, which captures known disturbances from ambient air temperature and the HVAC system’s supply air temperature. Note that from our previous studies, heat gains due to solar radiation are orders of magnitude less than those caused by other disturbances and inputs and hence, are not included here. Finally, \( f_{IG}(k) : \mathbb{N} \rightarrow \mathbb{R}^4 \) captures internal gains in each of the six zones on the 4th floor and consists of two terms:

\[
f_{IG}(k) = f_{IG}^c + f_{IG,m}(k) \quad \text{for } m \in \{f, w, s\},
\]

where \( f_{IG}^c \) is an unknown constant vector representing background heat gains due to idle appliances such as computers and printers. Functions \( f_{IG,m}(\cdot) \) for \( m \in \{f, w, s\} \), are unknown nonparametric functions that capture the time-varying heat gain due to occupancy, equipment and other unmodeled uncertainties such as reheating at the VAV boxes in fall, winter and spring, respectively. The system matrices \( A, B_u, B_{IG}, B_{ux}, B_{vu} \) and \( C \) are functions of tuning parameters: the window heat transmission coefficient \( (U_{win}) \), the convection coefficients of the interior wall (\( \gamma_{TW} \)), the exterior wall (\( \gamma_{EW} \)), the floor (\( \gamma_{floor} \)), and the ceiling (\( \gamma_{ceil} \)). Define \( \gamma := [U_{win}, \gamma_{TW}, \gamma_{EW}, \gamma_{floor}, \gamma_{ceil}] \) \( \in \mathbb{R}^{11} \), then to identify the physics-based model, we need to estimate the parameter vector \( \gamma \) as well as the functions \( f_{IG,m}(\cdot) \).

Next, we describe our approach for identifying this model.

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, \( \gamma \). Second, the nonparametric functions \( f_{IG,m}^{\nu}(\cdot) \) are estimated from the complete training dataset.

1) Parameter Estimation: For parameter estimation purposes, we first set \( f_{IG,m}^{\nu}(\cdot) = 0 \) during the weekend days, and evaluate them at a later point (Equations (16) and (17)). With \( f_{IG,m}^{\nu}(\cdot) = 0 \), (11) reduces to a purely parametric model:

\[
\begin{align*}
    x(k+1) &= Ax(k) + B_u v(k) + B_{IG} f_{IG}^c + \sum_{i=1}^{21} (B_{ux}, x(k) + B_{vu}, v(k)) u_i(k) \\
    y(k) &= C x(k).
\end{align*}
\]

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

\[
\begin{align*}
    \hat{\gamma} &= \arg \min_{\gamma > 0} \left( J_I + \sum_{m=1}^{2} \| y_m(k, \gamma) - \tilde{y}_m(k) \|_2^2 \right) \\
    \text{s.t.} \quad J_m &= \sum_{k=0}^{n_m} \| y_m(k, \gamma) - \tilde{y}_m(k) \|_2^2 \quad \text{for } m \in \{f, w, s\} \\
    y_m(k, \gamma) \text{ and } x_m(k, \gamma) \text{ satisfy (13)} \text{ with} \\
    x_m(0) &= x_{KF,m}(0) \\
    u_m(k) &= \tilde{u}_m(k), v_m(k) = \tilde{v}_m(k) \quad \forall \, k,
\end{align*}
\]

where \( \tilde{u}, \tilde{v} \) and \( \tilde{y} \) denote the measured inputs, disturbances, and zone temperatures, respectively. In other words, we choose \( \gamma \) 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 (the wall temperature for example is not), thus we estimate the initial states using a Kalman Filter \( x_{KF,m}(0) \), and set \( x_m(0) = x_{KF,m}(0) \). Furthermore, to compensate for the lack of sufficient excitation of the building, initial guesses for \( \gamma \) that are physically plausible are chosen. The optimal parameter values are similar to those reported in [18], hence are not included here due to space limitations.

2) Estimation of \( f_{IG,m}^{\nu}(\cdot) \) for Each Season: Let \( \mathcal{W}_m = \{1, 2, \ldots, n_m\} \) denote the set of weeks in the training data for season \( m \), and let \( f_{IG,m,w}^{\nu}(\cdot) \) be an instance of the internal gains function \( f_{IG,m}^{\nu}(\cdot) \) estimated for week \( w \) in \( \mathcal{W}_m \). The optimal \( f_{IG,m,w}^{\nu}(\cdot) \) is defined as the average of all estimates for a given season.

More specifically, let \( \hat{x}(k) \) and \( \tilde{y}(k) \) denote the predicted states and zone temperatures at time \( k \), with \( f_{IG,m,w}^{\nu}(k-1) = 0 \), i.e.,

\[
\begin{align*}
    \hat{x}(k) &= Ax(k-1) + B_u v(k-1) + B_{IG} f_{IG}^c + \sum_{i=1}^{21} (B_{ux}, x(k-1) + B_{vu}, v(k-1)) u_i(k-1) \\
    \tilde{y} &= C \hat{x}(k).
\end{align*}
\]

By noting \( x(k) = \hat{x}(k) + B_{IG} f_{IG,m,w}^{\nu}(k-1) \), we can estimate \( f_{IG,m,w}^{\nu}(k-1) \) by solving the following set of linear equations using Ordinary Least Squares:

\[
(CB_{IG}) \cdot f_{IG,m,w}^{\nu}(k-1) = \tilde{y}(k) - \hat{y}(k),
\]

where \( \hat{y}(k) \) represents measured zone temperatures at time \( k \). Finally, \( f_{IG,m}^{\nu}(\cdot) \) is chosen as the average of all the estimates:

\[
f_{IG,m}^{\nu}(k) = \frac{\sum_{w=1}^{n_m} f_{IG,m,w}^{\nu}(k)}{n_m} \quad \forall \, k.
\]
Therefore, the estimated function \( \hat{f}_{IG,m} \) takes into account the effect of hour of the day and day of the week on the internal gains.

C. Results

The identified model is tested on holdout test weeks from different seasons. The average daily prediction RMS errors by zone and season are reported in Table I. Figure 4 shows the estimated increase in zone temperatures due to internal gains for fall, winter and spring: \( B_{IG} \cdot (f_{IG} + \hat{f}_{IG,m}(k)) \) for \( m \in \{f, w, s\} \), respectively. Similar average internal gains are observed for all zones and seasons. The zones that correspond to open workspaces and conference rooms (“West”, “South”, “East” and “Center”) show discernible daily peaks in their internal gains profiles with a slight decrease during weekends. Furthermore, there is little variation in the internal gains profiles across different seasons.

V. MODEL COMPARISON

A. Prediction Accuracy

The high-dimensional physics-based model (Model B) is found to have a higher prediction accuracy compared to the low-dimensional data-driven model for the individual zones (Model A) presented in Section III-B: According to Table I, the mean RMS for Model B across zones is more than 0.1 degrees lower than for Model A. This is also illustrated in Figure 5, which shows 7-day open-loop predictions of the temperature of a selected holdout test week in the spring period, simulated with both models instantiated once with an initial condition that matches the measured temperatures. The increase in RMS from Model B to Model A is notably larger in the zones “East” (0.38) and “Center” (0.15), compared to the other zones (0.11, −0.03, 0.07, and 0.08).

B. Energy Efficient Control

In this section, we compare the performance of Model A and Model B for the purpose of energy efficiency. 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 \([20^\circ \text{C}, 22^\circ \text{C}]\) [29], and confining the control input to the physical limits of the HVAC system \([u_{\min}, u_{\max}]\).

This problem is formulated as follows:

\[
\begin{align*}
\min_{u, \varepsilon} & \quad \sum_{k=1}^{N} u(k) + \rho \|\varepsilon\|_2 \\
\text{s.t.} & \quad x(0) = \bar{x}(0) \\
& \quad x(k + 1) = \begin{cases} (9), & \text{Model A} \\
(11a), & \text{Model B} \end{cases} \\
& \quad u_{\min} - \varepsilon \leq u(k) \leq u_{\max} + \varepsilon \quad \forall k \in \{0, N - 1\} \\
& \quad T_{\min} \leq x(k) \leq T_{\max}, \quad \text{Model A} \\
& \quad T_{\min} \leq C x(k) \leq T_{\max}, \quad \text{Model B (11b)} \\
& \quad \forall k \in \{1, N\}
\end{align*}
\]

The temperature is initialized with the measured temperature \( \bar{x}(0) \) at the beginning of the week-long simulation. We use soft constraints on the control input with a penalty parameter \( \rho \) to ensure the feasibility of the problem. The penalty represents the cost of increasing the airflow beyond the operating limits (temporary shutdown or overuse, both of which are harmful to the system). To find the optimal control strategy, we make use of receding horizon control with a prediction horizon of three 15-minute time steps.

Figure 6 shows the temperature trajectory computed by the energy efficient controller (18) 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 \([20^\circ \text{C}, 22^\circ \text{C}]\), 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. An interesting observation is that the largest difference in the control strategies is detected in zones “East” and “Center”, which show a larger increase in RMS from
Model B to Model A. The dips of the computed control trajectory $u$ below the black dashed lines represent violations of the physical limits needed to maintain the temperature in the narrow range $[20^\circ, 22^\circ]$. 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 ($\approx 20$ 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, in essence, reduces the difference in the prediction quality by both controllers, 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. For temperature-critical zones in which precise temperature estimations are needed, however, one might still want to choose the fine-level Model B for analysis.

C. Qualitative Comparison

We summarize the findings in the following:

- Model A is more amenable to controller design due to its low dimensionality and hence considerably faster operation. This is of particular importance for scalability considerations, since the computation time grows exponentially with the number of state variables, which renders Model B computationally intractable for online operation beyond a certain complexity. Indeed we observe that the computation for one step of (18) exceeds 15 minutes — the discretization time — for a prediction horizon of five steps. Thus in frequency regulation, for instance, Model A’s low dimensionality makes it suitable for reserve determination which must be computed for a time horizon of 24 hours.

- Identifying Model A with semiparametric regression only relies on temperature and VAV airflow data and the physical adjacency of zones, in contrast to Model B, which requires knowledge of the BRCM toolbox and a large amount of geometry and construction data of the buildings, many of which are often unknown [18]. Hence, more effort is required to train the model on new buildings for Model B.

- The higher accuracy of Model B proves useful for applications such as the control of temperature-critical zones and evaluation of controller performance through simulations, whereas Model A is preferably used for controller design and in applications where less emphasis is put on estimation errors, e.g. at night when building occupancy is low.

- Model A assumes a uniform temperature among zones, which often encompass several rooms, whereas Model B can provide estimates for the temperature of individual rooms in a given zone. The number of parameters of Model A increases rapidly with the model complexity, which coupled with insufficient excitation of the system makes it hard to emulate the higher spatial granularity with Model A.

VI. CONCLUSION

We identified state-space models for the thermal behavior of SDH with semiparametric regression and a physics-based model. The internal gains due to occupants and electric devices were identified for different spatial granularities and different temporal seasons. We found the high-dimensional physics-based model to yield lower estimation errors than the low-dimensional data-driven model due to the inclusion of analytical temperature models based on physical parameters of the building, therefore allowing for higher granularity in temperature predictions. Under an energy efficient MPC
scheme, both models performed equally well, with the disadvantage of the physics-based model being computationally expensive due to its large number of states, which show an inherent bilinear relationship with inputs.

We note that the higher fidelity physics-based model should be used for controlling temperature-critical zones in buildings, since it provides higher granularity in addition to higher accuracy. The compact data-driven model, however, is a good alternative for devising a control strategy when less emphasis is put on estimation errors, e.g. at night when occupancy is low. In frequency regulation, the lower-dimensional data-driven model is more suitable for reserve determination as it requires planning over a longer time horizon, whereas the more accurate higher-dimensional physics-based model can be used in reserve provision to maintain the building temperature within comfort bounds and track the frequency regulation signal. Furthermore, while semi-parametric regression can be easily applied on any building with a modest requirement of recorded data, the physics-based model requires detailed geometry and construction data about the building, which in practice is often subject to large inaccuracies, and therefore hard to obtain.

We are currently designing a control scheme suitable for frequency regulation in commercial buildings, based on the findings outlined in this paper. Further, we will validate the identified models by implementing these control schemes 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] 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.

[7] E. Vrettos, F. Oldewurtel, F. Zha, and G. Andersson, “Robust Provision of Frequency Reserves by Office Building Aggregations,” Proceedings of the 19th IFAC World Congress, pp. 12068 – 12073, 2014.

[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. Shahbakti, and A. Sangiovanni-Vincentelli, “Handling Model Uncertainty in Model Predictive Control for Energy Efficient Buildings,” Energy and Buildings, 2014.

[10] B. Sun, P. Luh, Q. Jia, Z. Jiang, F. Wang, and C. Song, “Building Energy Management: Integrated Control of Active and Passive Heating, Cooling, Lighting, Shading, and Ventilation Systems,” IEEE Transactions on Automation Science and Engineering, 2012.

[11] D. Sturzenegger, D. Gyavalis, M. Morari, and R. Smith, “Semi-Automated Modular Modeling of Buildings for Model Predictive Control,” BuildSys 2012 – Workshop of SCM SensSys Conference, 2012.

[12] 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.

[13] M. Duffy, M. Hiller, D. Bradley, W. Keilholz, and J. Thornton, “TRNSYS - Features and Functionality for Building Simulation,” IBPSA Conference, pp. 1950 – 1954, 2009.

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

[15] A. Aswani, N. Master, J. Taneja, V. Smith, A. Krioukov, D. Culler, and C. Tomlin, “Identifying Models of HVAC Systems Using Semi-parametric Regression,” American Control Conference, 2012.

[16] Y. Ma, F. Borrelli, B. Hencye, and B. Coffey, “Model Predictive Control for the Operation of Building Cooling Systems,” IEEE Transactions on Power Systems, vol. 20, pp. 796–803, 2012.

[17] Y. Ma, G. Anderson, and F. Borrelli, “A Distributed Predictive Control Approach to Building Temperature Regulation,” American Control Conference, pp. 2089–2094, 2011.

[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] “Commercial Buildings Energy Consumption Survey (CBECS): Overview of Commercial Buildings, 2003,” Energy Information Administration, U.S. Department of Energy, Tech. Rep., 2008. [Online]. Available: http://www.eia.doe.gov/emeu/cbeecs/cbeecs2003/overview1.html

[20] “sMAP: The Simple Measurement and Actuation Profile.” [Online]. Available: http://www.cs.berkeley.edu/~stevedh/smaph2/

[21] S. Dawson-Haggerty, A. Krioukov, and D. E. Culler, “Experiences Integrating Building Data with sMAP,” University of California, Berkeley, Tech. Rep., 2012.

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

[23] “PV Performance Modeling Collaborative.” [Online]. Available: https://pvpmc.sandia.gov/

[24] Y. Lin, T. Middelkoop, and P. Barooah, “Issues in Identification of Control-Oriented Thermal Models of Zones in Multi-Zone Buildings,” 51st IEEE Conference on Decision and Control, December 2012.

[25] F. Oldewurtel, A. Parisio, C. N. Jones, M. Morari, D. Gyavalis, M. Gwerder, V. Staub, B. Lehmann, and K. Wirth, “Energy Efficient Building Climate Control Using Stochastic Model Predictive Control and Weather Predictions,” American Control Conference, 2010.

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

[27] W. Hardle, H. Liang, and J. Guo, Partially Linear Models. Springer, 2000.

[28] S. Koehler and F. Borrelli, “Building Temperature Distributed Control via Explicit MPC and “Trim and Respond” Methods,” European Control Conference, 2013.

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