Building Model Identification during Regular Operation
– Empirical Results and Challenges

Qie Hu, Frauke Oldewurtel, Maximilian Balandat, Evangelos Vrettos, Datong Zhou, and Claire J. Tomlin

Abstract— The inter-temporal consumption flexibility of commercial buildings can be harnessed to improve the energy efficiency of buildings, or to provide ancillary service to the power grid. To do so, a predictive model of the building's thermal dynamics is required. In this paper, we identify a physics-based model of a multi-purpose commercial building including its heating, ventilation and air conditioning system during regular operation. We present our empirical results and show that large uncertainties in internal heat gains, due to occupancy and equipment, present several challenges in utilizing the building model for long-term prediction. In addition, we show that by learning these uncertain loads online and dynamically updating the building model, prediction accuracy is improved significantly.

I. INTRODUCTION

Commercial buildings account for more than 35% of electricity consumption in the U.S., 39% of which is due to heating, ventilation and air conditioning (HVAC) systems [1]. Energy consumption of HVAC systems can be partly shifted in time without compromising occupant comfort, because of buildings' inherent thermal capacity. As a result, there has been extensive research, using frameworks such as model predictive control (MPC), trying to harness this inter-temporal consumption flexibility and minimize energy usage of buildings [2], [3]. More recently, the feasibility of using buildings to provide ancillary services, such as frequency regulation, to the power grid has also been studied [4], [5], [6], [7], [8]. Such applications require accurate models describing the thermal dynamic behavior of the buildings.

A. Desired Model Features and Challenges

The building models should be identified using actual experimental data and capture realistic disturbances, such as internal gains. Furthermore, bilinear, multi-zone models may quantify buildings' electricity consumption flexibility more precisely: the bilinear thermal dynamics naturally arise from the physics of the HVAC system (refer to Section III for details); controllers designed using multi-zone models can allow the temperature of a room to fluctuate when it is unoccupied, for instance, hence achieving energy savings which

are not possible when simplified models that approximate the building as a single zone are used.

However, there are many challenges in identifying such a building model. First, actual buildings often have different types of spaces that are subject to very different uncertainties, e.g. occupancy, which are difficult to capture. Second, buildings are often not sufficiently excited, as they must satisfy strict regulatory requirements during regular operation, which limit the type and duration of excitation experiments that can be conducted [9].

To circumvent these difficulties, various approaches have been taken in the research community. In [9], [10], [11], the authors identify data-driven linear models for a single type of building space. Lin *et al.* [12] conduct frequency regulation experiments under a controlled environment without disturbances such as solar radiation and occupants, thus simplifying the building model required to design the controller. Faced with insufficient excitation of buildings, Mehdi *et al.* [13] reduce the required model complexity by carrying out their experiments in a single room, whereas others [14], [15], [16] use lumped thermal models that approximate the building as a single zone, with an average building temperature. Finally, the authors of [17], [18] identify and test multi-zone models for a single floor and an entire building, respectively, using simulated data where uncertainties are precisely controlled or removed, and arbitrary excitations can be simulated.

These approaches are valuable in providing estimates of a building's consumption flexibility, however none of them delivers a model that satisfies all the aforementioned desired features for a more precise quantification of the building's flexibility.

B. Contributions

In this paper, we identify a semi-parametric model for a multi-zone commercial building during regular operation. Our main contributions are the following:

- We propose a procedure to identify a physics-based model of a multi-zone building, that is easy to implement with the building in regular operation, and captures internal gains such as occupancy. This procedure uses excitation experiments that actively perturb the building and generate data that can be used for more accurate parameter identification.
- We provide an analysis of the model's prediction accuracy versus its prediction horizon, and show that this model is limited in making long-term predictions, partly

Q. Hu, F. Oldewurtel, M. Balandat, D. Zhou and C.J. Tomlin are with the Department of EECS, University of California, Berkeley. {qiehu,oldewurtel,balandat,datong.zhou,tomlin}@berkeley.edu. E. Vrettos is with Power Systems Laboratory, ETH Zürich. vrettos@eeh.ee.ethz.ch.
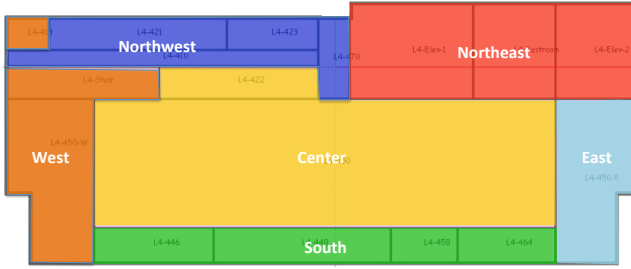
Fig. 1. Zones for the 4th floor of Sutardja Dai Hall (SDH).

due to large disturbances such as internal gains, which are uncertain.

- Finally, we propose to dynamically update the estimate of internal gains based on current temperature measurements, and show that this significantly improves the building model's prediction accuracy.

This paper is organized as follows: We begin by describing the building and the excitation experiments in Section II. We then present the building model in Section III. Section IV describes our identification method for a fixed model and reports the prediction results. A model that is updated online is presented in Section V. Finally, we provide a discussion of uses and challenges related to building models.

Notation: Unless stated otherwise, subscripts in italics as in u_i denote instances of variables. Upright subscripts as in U_{win} denote variable names.

II. BUILDING AND EXCITATION EXPERIMENTS

A. Building

Sutardja Dai Hall (SDH) is a building located on the University of California, Berkeley campus. For ease of presentation, we focus on the entire 4th floor of SDH, which has a total floor area of 1300m². As shown in Figure 1, we aggregate the rooms into 6 zones: Northwest (NW), West (W), South (S), East (E), Northeast (NE) and Center (C). The north-side rooms are grouped into 2 zones because of their distinct characteristics: rooms in zone Northwest are offices with windows, whereas zone Northeast includes elevators, restrooms and utility rooms, and does not have windows.

This 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 that drive air over cooling coils, cooling it down to a desired supply air temperature, and then distribute the air to VAV boxes that govern the airflow to different building zones. The supply air may be reheated at the VAV box before entering the room. The 4th floor of SDH is served by 21 VAV boxes.

B. Excitation Experiments and Data

Data was collected during 11 non-consecutive weeks between September 2014 and June 2015. This time span includes periods when the building was under normal operation as well as periods with excitation experiments. Recorded data points include room temperatures measured at all VAV boxes on the 4th floor of SDH, air inflow rates from each VAV box, HVAC system's supply air temperature, outside ambient air temperature and solar radiation data recorded from a nearby weather station [20].

For accurate parameter identification, temperatures of neighboring zones should not be strongly correlated [21]. For buildings in regular operation, this is generally achievable through forced response experiments. Because of commercial buildings' large thermal inertia, each forced excitation must last sufficiently long before temperature changes are observed. With these points in mind, we conducted our experiment as follows: Starting at 8am, every 2 hours, the supply airflow rate to one zone is set to its maximum value, minimum airflow rates are set for each of its neighboring zones and a random airflow rate is chosen for each remaining zone. This is repeated for each of the 6 zones. This experiment is performed during weekends as (a) it minimizes effects due to building occupancy on our data, and thus the subsequent parameter identification; (b) temporary violation of comfort constraints during the weekend was allowed.

III. BUILDING MODEL

A. RC Modeling and the BRCM Toolbox

We derive a Resistance-Capacitance (RC) model for our building using the Building Resistance-Capacitance Modeling (BRCM) MATLAB toolbox [17]. The RC modeling methodology first decomposes a building into building elements (BE), such as the bulk volume of air in each room, walls, floors and ceilings. Then, an electric analogy is used to derive an equivalent electrical circuit whose resistances and capacitances represent thermal resistances and thermal capacitances of the BEs, and voltages and currents represent temperatures of BEs and heat transfers between those. The thermal resistances and capacitances of BEs are completely characterized by their geometry and construction data such as density, convection coefficient and specific heat capacity.

In the BRCM toolbox, a building model consists of two parts: a thermal submodel and external heat flux submodels (EHFM). The thermal submodel describes passive heat transfer between the BEs and the EHFMs capture heat gain or loss due to external inputs and disturbances such as the outside environment. The BRCM toolbox semi-automates the derivation of an RC model by automatically computing the thermal submodel using the electrical circuit analogy and an input file that contains geometry and construction data of all BEs (e.g. we use an EnergyPlus file developed for the 4th floor of SDH as our input file). The EHFMs can be user defined, as different buildings may be subject to distinct inputs and disturbances.

B. Building Model

In this section, we first describe the EHFMs for our building and then present the final state-space model of the building.

There are 3 EHFMs:

- Building hull: convective heat transfer and solar radiation gains across exterior walls and windows.
- HVAC: heat gain from the HVAC system.
- Internal gains (IG): heat gain due to occupancy, electrical appliances and other unmodeled disturbances.

Let $x \in \mathbb{R}^{289}$ be the state vector whose elements are the temperatures of all BEs on the 4th floor of SDH[1], $u \in \mathbb{R}^{21}$ be the air inflow rate from the 21 VAV boxes on this floor and $v := \begin{bmatrix} v_{T_a}; v_{T_s}; v_{solE}; v_{solS}; v_{solW}; v_{solN} \end{bmatrix} \in \mathbb{R}^6$ be the disturbance vector whose elements represent ambient air temperature, supply air temperature from the HVAC system and solar radiation from the four geographical directions, respectively.

Building Hull: If the i-th BE is connected to the building hull, e.g. a room adjacent to the exterior wall, then the external heat fluxes acting on it due to the outside environment are modeled as:

$$
\begin{aligned}
q_{\text{BH},i} = {} & \gamma_{\text{EW}}\, a_{\text{EW},i} \big(v_{T_a}(t) - x_i(t)\big) + \gamma_{\text{absorp}}\, a_{\text{EW},i}\, v_{\text{sol},i}(t) \\
& + U_{\text{win}}\, a_{\text{win},i} \big(v_{T_a}(t) - x_i(t)\big) \\
& + \gamma_{\text{winSolAbs}}\, a_{\text{win},i}\, v_{\text{sol},i}(t),
\end{aligned}
\tag{1}
$$

where $a_{\text{EW},i}$ and $a_{\text{win},i}$ are the total areas of the exterior wall and the window respectively, and $v_{\text{sol},i} \in \{v_{\text{solE}}; v_{\text{solS}}; v_{\text{solW}}; v_{\text{solN}}\}$ is the solar radiation affecting this BE. γ_{EW}, γ_{absorp}, U_{win} and $\gamma_{\text{winSolAbs}}$ are tuning parameters of the model and their descriptions are given in Table I.

HVAC: If the i-th BE is a room equipped with at least one VAV box, then the heat flux acting on it is:

$$
q_{\text{HVAC},i} = c_p \Big(\sum_{j \in \mathcal{B}_i} u_j(t) \Big) \cdot \big(v_{T_s}(t) - x_i(t)\big), \tag{2}
$$

where c_p is the specific heat capacity of air, $\mathcal{B}_i$ is the set of VAV boxes serving the i-th room and u_j is the j-th element of u. Note that due to the lack of temperature measurements of the supply air at the outlet of VAV boxes, v_{T_s} is the supply air temperature upstream of the VAV boxes' heating coils, i.e., heat gains due to reheating at the VAV boxes are not modeled by (2), but are captured by the internal gains EHFM in our model.

Internal Gains: If the i-th BE is a room, then it is also subject to internal gains, which are modeled as:

$$
q_{\text{IG},i} = a_{\text{floor},i} \big(c_{\text{IG},i} + f_{\text{IG},i}(t)\big), \tag{3}
$$

where $a_{\text{floor},i}$ is the room's floor area. $c_{\text{IG}} \in \mathbb{R}^6$ is an unknown constant vector representing a background time-invariant heat gain per unit area in each of the 6 zones, due to idle appliances. The function $f_{\text{IG}}(t) : \mathbb{R} \to \mathbb{R}^6$ is an unknown function that captures time-varying internal gains in different zones. Finally, $c_{\text{IG},i}$ and $f_{\text{IG},i}(t)$ are the relevant elements of c_{IG} and $f_{\text{IG}}(t)$ that correspond to the i-th room.

[1]In the EnergyPlus file for the 4th floor of SDH, each wall, floor and ceiling is decomposed into 2 to 3 BEs. In RC building models, the temperature of each BE is modeled by one state variable, thus the model of the 4th floor of SDH has a large number of states: 289.

TABLE I
MODEL PARAMETERS.

Parameter	Description	Value [Unit]
γ_{EW}	exterior wall convection coeff.	10.5 [W/(m^{2}K)]
γ_{IW}	interior wall convection coeff.	29.4 [W/(m^{2}K)]
γ_{floor}	floor convection coeff.	51.5 [W/(m^{2}K)]
γ_{ceil}	ceiling convection coeff.	44.3 [W/(m^{2}K)]
γ_{absorp}	ext. wall solar absorption coeff.	0.75 [-]
$\gamma_{\text{winSolAbs}}$	window solar absorption coeff.	0.03 [-]
U_{win}	window heat transmission coeff.	0.63 [W/(m^{2}K)]
$c_{\text{IG,NW}}$	background heat gain in zone NW	0.3 [W/m^2]
$c_{\text{IG,W}}$	background heat gain in zone W	8.0 [W/m^2]
$c_{\text{IG,S}}$	background heat gain in zone S	18.8 [W/m^2]
$c_{\text{IG,E}}$	background heat gain in zone E	8.0 [W/m^2]
$c_{\text{IG,NE}}$	background heat gain in zone NE	11.0 [W/m^2]
$c_{\text{IG,C}}$	background heat gain in zone C	8.0 [W/m^2]

After defining all EHFMs, the BRCM toolbox automatically generates the following model:

$$
\begin{aligned}
\dot{x}(t) = {} & A_t x(t) + B_t \big[q_{\text{BH}}\big(x(t), v(t)\big) \\
& + q_{\text{HVAC}}\big(x(t), v(t), u(t)\big) + q_{\text{IG}}(t) \big], \\
= {} & A x(t) + B_v v(t) + B_{\text{IG}}\big(c_{\text{IG}} + f_{\text{IG}}(t)\big) \\
& + \sum_{j=1}^{21} \big(B_{xu_j} x(t) + B_{vu_j} v(t)\big) u_j(t)
\end{aligned}
\tag{4}
$$

where in the first equality, A_t, B_t represent the thermal submodel, and $q_{\text{BH}}(\cdot)$, $q_{\text{HVAC}}(\cdot)$, $q_{\text{IG}}(\cdot)$ are the EHFMs, which are vector-valued functions as follows: if the i-th BE is not subject to a specific EHFM, say HVAC, then $q_{\text{HVAC},i} = 0$, otherwise $q_{\text{HVAC},i}$ is given by Equation (2). The second equality is obtained by expressing q_{BH}, q_{HVAC} and q_{IG} as functions of x, u and v, using (1) to (3). The bilinearities in (4) naturally arise from the physics of the HVAC system (refer to Equation (2)).

Finally, we discretized the model using a fixed time step of 15 min to obtain its approximate discrete time model, which is semi-parametric and bilinear:

$$
\begin{aligned}
x(k+1) = {} & A x(k) + B_v v(k) + B_{\text{IG}}\big(c_{\text{IG}} + f_{\text{IG}}(k)\big) \\
& + \sum_{j=1}^{21} \big(B_{xu_j} x(k) + B_{vu_j} v(k)\big) u_j(k) \\
y(k) = {} & C x(k),
\end{aligned}
\tag{5}
$$

where x, u and v are as defined before, and $y \in \mathbb{R}^6$ represents the measured average temperature of each zone.

IV. MODEL IDENTIFICATION

The model identification process is carried out in two steps. In Section IV-A, we identify the model parameters listed in Table I. To simplify the parameter identification process, we use the approximation $f_{\text{IG}}(k) = 0$ during weekend days, in order to reduce (5) to a purely parametric model. With the optimal parameter values in hand, we then identify the function $f_{\text{IG}}(\cdot)$ in Sections IV-B.

A. Parameter Identification

The model parameters are estimated using data collected during two weekends in spring and summer, when excitation experiments were carried out. The identified model is then validated on data collected during a weekend in fall (using $f_{\text{IG}}(k) = 0$).

	NW	W	S	E	NE	C	Mean
Training	0.46	0.41	0.24	0.34	0.27	0.28	0.333
Validation	0.62	0.57	0.31	0.28	0.39	0.31	0.413

Let $\gamma \in \mathbb{R}^{13}$ be the parameter vector whose elements are those parameters listed in Table I. We choose the optimal $\hat{\gamma}$ that solves the following optimization problem:

$$\hat{\gamma} = \arg\min \ \sum_k \|y(k,\gamma) - \bar{y}(k)\|^2$$
$$\text{s.t. } y(k,\gamma) \text{ and } x(k,\gamma) \text{ satisfy (5) with } f_{\text{IG}}(k) = 0 \ \forall k$$
$$u(k) = \bar{u}(k), \ v(k) = \bar{v}(k) \ \forall k$$
$$x(0) = x_{\text{KF}}(0)$$
$$\gamma > 0, \ 0 \leq \gamma_{\text{absorp}}, \gamma_{\text{winSolAbs}} \leq 1, \tag{6}$$

where $\bar{y}$, $\bar{u}$ and $\bar{v}$ are measured zone temperatures, inputs and disturbances, respectively; and $x_{\text{KF}}(0)$ is the initial state estimated using a Kalman Filter. Estimating $x(0)$ is necessary since not all states can be measured (for example, the wall temperatures are not). In other words, we choose the set of parameters such that when the model is simulated with this set of parameters and the measured inputs and disturbances, the sum of squared errors between the measured temperatures and the simulated temperatures is minimized. We can use prior knowledge about the building to compensate for limited excitation of the system, e.g. we use initial guesses for parameter values that are physically plausible.

The identified parameter values are reported in Table I. The root-mean-square (RMS) errors between the model's simulated temperature for different zones and the measured temperatures for the training data are shown in Table II. This table also shows the RMS prediction errors when the identified model is used to predict a validation dataset (also on weekend data, using $f_{\text{IG}}(k) = 0$).

B. Identification of the Time-Varying Internal Gains

A random subset of 8 weeks is selected from the entire dataset and used as training data for estimating the time-varying internal gains function $f_{\text{IG}}(\cdot)$, and the remaining 3 weeks of data are used as a validation set. For each week w in the training set, we estimate an instance of this function, $f_{\text{IG},w}(\cdot)$. The final estimate of the function $f_{\text{IG}}(\cdot)$ is defined as the average of all estimates $f_{\text{IG},w}(\cdot)$.

More specifically, at each time k, let $\tilde{x}_w(k)$ and $\tilde{y}_w(k)$ denote the simulated state and measurement vectors with $f_{\text{IG},w}(k-1) = 0$, i.e.,:

$$\tilde{x}_w(k) = A x_w(k-1) + B_v v_w(k-1) + B_{\text{IG}} c_{\text{IG}}$$
$$+ \sum_{j=1}^{21} \left(B_{xu_j} x_w(k-1) + B_{vu_j} v_w(k-1) \right)$$
$$\cdot u_{w,j}(k-1)$$
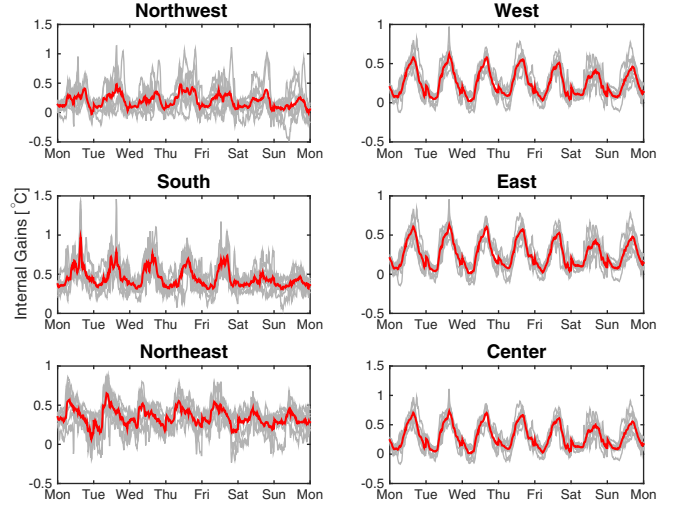$$\tilde{y}_w = C\tilde{x}_w(k). \tag{7}$$

Fig. 2. Estimated increase in temperature due to internal gains by zone. Gray lines are estimates for individual weeks in the training dataset, red lines are their averages.

By noting $x_w(k) = \tilde{x}_w(k) + B_{\text{IG}} f_{\text{IG},w}(k-1)$, we can estimate $f_{\text{IG},w}(k-1)$ by solving the following set of linear equations using ordinary least-squares:

$$(CB_{\text{IG}}) \cdot f_{\text{IG},w}(k-1) = \bar{y}_w(k) - \tilde{y}_w(k), \tag{8}$$

where $\bar{y}_w(k)$ is the measured zone temperature at time k from the w-th training week. Finally, the estimate $\hat{f}_{\text{IG}}(\cdot)$ is obtained by:

$$\hat{f}_{\text{IG}}(k) = \frac{1}{8} \sum_{w=1}^{8} f_{\text{IG},w}(k) \quad \text{for all } k. \tag{9}$$

C. Impact of Internal Gains

The estimated average increase in room temperature due to internal gains, i.e., $B_{\text{IG}}\big(c_{\text{IG}} + \hat{f}_{\text{IG}}(k)\big)$ is shown in red in Figure 2. It can be observed that internal gains profiles vary on both long and short time horizons from approximately 0°C to 1°C. A slightly larger increase in temperature of approximately 1.4°C is reported in [9] for a similar office space. This may be because the internal gains term in their model also includes heat gain from solar radiation, whereas our model captures the effects of solar radiation separately. Observe that the internal gains profiles increase during the day, peaking in the early afternoon and then slowly decrease, reaching a minimum at night. In addition, the profiles' peaks are slightly lower during the weekends. These patterns coincide with when building occupants typically come into and leave the office.

The gray lines in Figure 2 show the same quantity estimated for each training week, i.e., $B_{\text{IG}}\big(c_{\text{IG}} + f_{\text{IG},w}(k)\big)$. It can be observed that different zones experience different variations in internal gains across the training weeks. The zones West, East and Center are workspaces of students who tend to have regular schedules and hence, more regular internal gains patterns. On the other hand, the remaining three zones experience more uncertainty in internal gains, possibly due to the presence of windows, elevators and
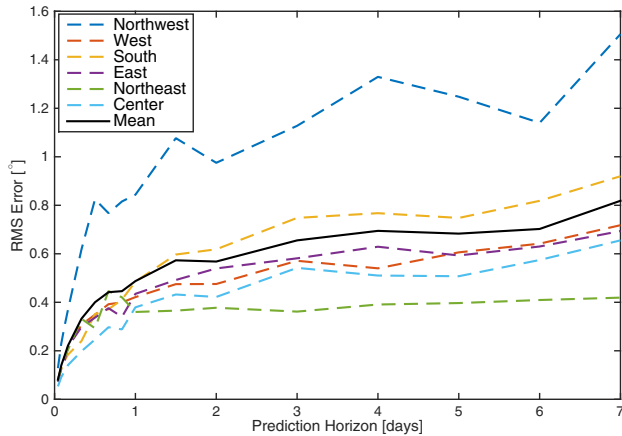
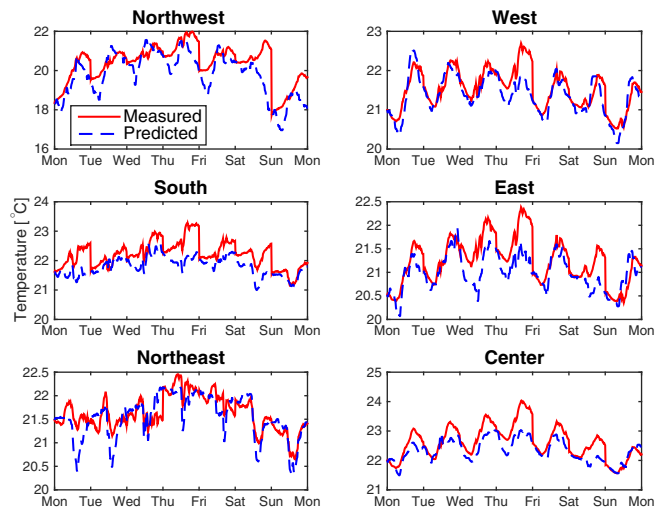Fig. 3. RMS prediction error versus prediction horizon.

Fig. 4. Predicted temperature using fixed internal gains function, by zone.

Fig. 5. Predicted temperature using online updated internal gains function, by zone.

staircases, and known inaccuracies in the EnergyPlus input file for zone Northwest.

The identified model with $\hat{f}_{IG}$ (a fixed function) is used to make 24-hour predictions of zone temperatures for all 3 weeks in the validation set, i.e., the state vector is estimated by a Kalman Filter every 24 hours. Figure 4 shows the results for one of these weeks and the average RMS prediction error for all validation weeks is reported in Table III. Furthermore, Figure 3 shows that the model's prediction accuracy decreases with increasing prediction horizon, which could be explained by uncertainties in internal gains as well as model inaccuracies. It is interesting to observe that the RMS prediction error for zone Northwest is the largest and it also increases the fastest as the prediction horizon increases, which is in accordance with this zone experiencing large variations in internal gains (Figure 2) and its geometry data in the EnergyPlus input file being inaccurate.

V. ONLINE UPDATE OF THE INTERNAL GAINS FUNCTION

The previous section shows that the large time-varying internal gains are difficult to capture *a priori*, nevertheless, they can significantly affect our model's prediction accuracy. In light of this, we consider a learning based approach in this section, where we update the internal gains function $f_{IG}(\cdot)$ online using past observations.

In other words, instead of estimating a fixed function $\hat{f}_{IG}(\cdot)$ *a priori*, at every time k, we estimate $f_{IG}(k-1)$ using (7) and (8), and then use the following simple model

$$f_{IG}(k) = f_{IG}(k-1) \qquad (10)$$

to obtain an online estimate of $f_{IG}(k)$, which is then used in (5) to predict $x(k+1)$ and $y(k+1)$. The intuition for (10) is that internal gains do not change significantly from time $k-1$ to k (i.e., 15 min).

This model is simulated for all 11 weeks of data, one of which is shown in Figure 5. The average RMS prediction errors are reported in Table III. Thus, by dynamically updating $f_{IG}(\cdot)$ online, the model's prediction accuracy is improved by 36% on average compared with when a fixed internal gains function was used.
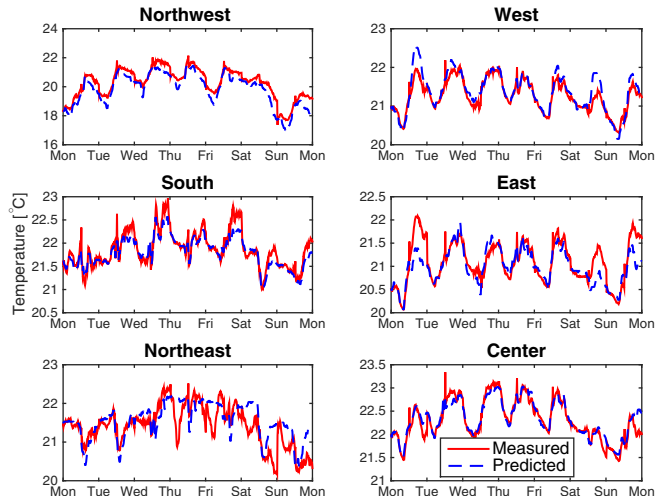
VI. DISCUSSION

In Section IV, we conducted excitation experiments to actively perturb our building. Data collected during the experiments and additional weekends is used with the approximation that the time-varying internal gains are zero, to identify the model parameters. Then, we estimate a fixed internal gains function, $f_{IG}(\cdot)$, using 8 weeks of measurements. The resulting model is used to make 24-hour predictions of the building's temperature profile for 3 additional weeks, and an average RMS error of $0.48°C$ is achieved. Figures 2 and 3 suggest that our building is subject to large uncertain internal gains which make accurate long term predictions difficult. For buildings that are subject to fewer uncertainties, a model identified using this procedure may achieve better prediction accuracy.

In addition, there are several approaches that can be taken to further enhance the model's prediction accuracy. When weekend data is used to identify the model parameters, more sophisticated approximations of the occupancy function, such

TABLE III
RMS PREDICTION ERROR BY ZONE FOR A MODEL WITH FIXED f_{IG} AND
USING ONLINE UPDATED f_{IG}.

	NW	W	S	E	NE	C	Mean
Fixed f_{IG}	0.84	0.42	0.48	0.43	0.36	0.38	0.485
Online Updated f_{IG}	0.50	0.31	0.15	0.41	0.32	0.16	0.308

as sinusoids, can be used. Sensors can be installed in the VAV boxes to measure the temperature of the supply air, from the HVAC system, downstream of the heating coils. Moreover, occupancy sensors can be used to improve the estimate of internal gains and hence the model's prediction quality.

In Section V, we dynamically updated our estimate of the internal gains function using current temperature measurements. More specifically, we assume the current heating load, due to internal gains, remains constant during the next time step. We apply this model to make 24-hr predictions and demonstrate that its prediction accuracy is significantly improved (compare Figure 4 with Figure 5). In addition, using more sophisticated regression techniques and taking into account other factors such as past heating loads and room temperatures may further improve the prediction accuracy and extend the prediction horizon.

For the frequency regulation application, a model is first applied to estimate the building's power consumption for the next 24 hours, in order to determine its reserve capacity for the day-ahead reserve market. MPC can then be used to provide these reserves without violating comfort constraints. Thus, by learning $f_{IG}(\cdot)$ online and dynamically updating it in an MPC controller, errors from the 24-hour prediction may be corrected during reserve provision.

VII. CONCLUSIONS

We describe an approach to construct a physics-based model of a multi-zone commercial building, which uses experimental data measured during regular building operation. We show that large uncertainties in internal gains present several challenges in applying the model for long term prediction of a building's thermal dynamics. In addition, we show that by dynamically updating the estimates of internal gains, the model's prediction accuracy is improved significantly.

Future work will investigate the trade-off between uncertainty in internal gains and prediction accuracy, as well as the necessary model complexity for a good control performance, in particular for harnessing building flexibility. In addition, we are working on experimentally verifying the performance of controllers designed using our building model.

ACKNOWLEDGMENT

REFERENCES

[1] "The annual energy outlook 2013," US Energy Information Administration, Tech. Rep., 2013.

[2] 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.

[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] F. Baccino, F. Conte, S. Massucco, F. Silvestro, and S. Grillo, "Frequency regulation by management of building cooling systems through model predictive control," *18th power systems computation conference*, pp. 1–7, August 2014.

[5] M. Maasoumy, C. Rosenberg, A. Sangiovanni-Vincentelli, and D. Callaway, "Model predictive control approach to online computation of demand-side flexibility of commercial buidings HVAC systems for supply following," *American Control Conference*, pp. 1082–1089, June 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*, 2014.

[7] E. Vrettos, F. Oldewurtel, and G. Andersson, "Robust energy-constrained frequency reserves from aggregations of commercial buildings," *IEEE Transactions on Power Systems*, 2016.

[8] 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*, September 2014.

[9] 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.

[10] A. Aswani, N. Master, J. Taneja, A. Krioukov, D. Culler, and C. Tomlin, "Energy-efficient building HVAC control using hybrid system LBMPC," *4th IFAC Nonlinear Model Predictive Control Conference*, August 2012.

[11] D. Zhou, Q. Hu, and C. J. Tomlin, "Model comparison of a data-driven and a physical model for simulating HVAC systems," *arXiv:1603.05951*, 2016.

[12] Y. Lin, P. Barooah, S. Meyn, and T. Middelkoop, "Experimental evaluation of frequency regulation from commercial building HVAC system," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 776–783, 2015.

[13] M. Maasoumy, M. Razmara, M. Shahbakhti, and A. Sangiovanni-Vincentelli, "Handling model uncertainty in model predictive control for energy efficient buildings," *Energy and Buildings*, vol. 77, pp. 377–392, 2014.

[14] F. Oldewurtel, A. Parisio, C. Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and K. Wirth, "Energy efficient building climate control using stochastic model predictive control and weather predictions," *American Control Conference*, pp. 5100–5105, July 2010.

[15] Y. Ma, F. Borrelli, B. Hencey, B. Coffey, S. Bengea, and P. Haves, "Model predictive control for the operation of building cooling systems," *IEEE Transactions on control systems technology*, pp. 796–803, 2011.

[16] H. Hao, T. Middelkoop, P. Barooah, and S. Meyn, "How demand response from commercial buildings will provide the regulation needs of the grid," *50th Annual Allerton Conference on communication, control and computing*, pp. 1908–1913, October 2012.

[17] 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.

[18] 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*, pp. 588–602, 2012.

[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/cbecs/cbecs2003/overview1.html

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

[21] 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.