Identification of control-oriented thermal models of rooms in multi-room buildings

Yashen Lin [1*], Timothy Middelkoop [2], Prabir Barooah [1]

[1] *Department of Mechanical and Aerospace Engineering, University of Florida*

[2] *Department of Industrial and Manufacturing Systems Engineering, University of Missouri*

Abstract

Model-based control for improving energy efficiency of buildings has been a popular topic of late. Smart control requires a predictive model of the building's thermal dynamics. Due to the complexity of the underlying physical processes, usually system identification techniques are used to identify parameters of a physics-based grey-box model. We investigate questions of required model structure and identification techniques for parameter estimation of a single zone model through a combination of analysis and experiments. Our results indicate that a second-order model can reproduce the input-output behavior of a full-scale model with 13 states. We also show that data collected during usual operation leads to poor parameter estimates that may nevertheless appear to predict the temperature well. The error becomes apparent when there is sufficient difference among various inputs and the output. We propose an algorithm to overcome these issues that involve specific forced-response tests. The results of this investigation are expected to provide guidelines on do's and don't's in modeling and identification of buildings for control.

1. Introduction

Model-based control of HVAC (Heating, Ventilation, and Air Conditioning) systems has generated excitement in the community of control researchers in re-

* Corresponding author. Email address: yashenlin@ufl.edu. Office tel: 352-392-7540.

This work has been supported by the National Science Foundation by Grants CNS-0931885 and ECCS-0955023.

cent years. This problem is of great societal importance since buildings account for 34% of total energy use in the United States and HVAC account for roughly half of that [1]. If novel control techniques can increase the efficiency of HVAC systems, it will lead to substantial savings in total energy use. A popular candidate for control of building HVAC systems is MPC (Model Predictive Control); a number of recent papers have been published on this topic; see [2, 3, 4] and references therein.

Demand for heating, cooling and ventilation air varies from room to room, and the control algorithms need to be targeted to each room based on its state. MPC requires a model that can be used to predict the evolution of temperature and humidity (and possibly other environmental variables such as CO_2 concentration) of a room given the inputs, which include mass flow rate and temperature of the supply air, temperatures of the surroundings, and heat gains. The underlying processes that govern the dynamics of temperature evolution are complex and uncertain, so simplified lumped parameter models need to be used. A question that is relevant to control designers is, given a prescribed degree of accuracy, how to choose model structure and identify its parameters that achieve that degree of accuracy? This paper attempts to answer this question. We limit ourselves to modeling a single room in a building with possibly many rooms.

The lumped parameter models that researchers typically use to model temperature dynamics of a room—and sometimes of an entire building—is a combination of R-C network (resistor-capacitor) models that capture heat transfer through solid surfaces. The model has a non-linear term that captures the effect of enthalpy exchange with the outside due to the supply and exhaust air. We restrict ourselves to this class of *non-linear R-C network models*, which is denoted by $\mathscr{M}_{RC}(n,q,p)$, where n refers to state dimension, q to the number of inputs, and p to the number of uncertain parameters. Two questions are relevant. Q1 (Minimal model complexity question): What is the minimum model complexity (measured by n,q,p) that is required to predict the temperature dynamics of a single room with an acceptable degree of accuracy? In this paper, we take "acceptable degree of accuracy" to mean a peak error in temperature prediction of less than 3°F. The reason is that the temperature distribution inside a room may have spatial variation of higher than 3°F. ASHRAE (American Society of Heating, Refrigeration and Air-conditioning Engineers) standard 55 [5] mandates that vertical temperature stratification in an occupied zone should be less than 5.4°F, meaning that this degree of spatial variation is expected. Q2 (Model calibration question): How to identify the values of the uncertain parameters (typically, resistances and capacitances) from measured data, and what kind of data are required, to achieve this

level of accuracy?

It should be noted that there is a long history of using R-C networks for modeling building thermal dynamics; see the review article [6]. Historically, they first arise in deriving lumped models of heat transfer through solid surfaces by discretizing the heat equation. However, whether they are fundamentally capable of reproducing observed data with a prescribed degree of accuracy for a room or a building has not been examined. It has been merely accepted that they are. In fact, ASHRAE (American Society of Heating, Refrigeration and Air conditioning Engineers) handbook [7] describes how to determine the R-C values for a solid surface given its material and construction type. This information is now available in software such as HAP [8]. The resistances and capacitances of a R-C network model are carefully chosen to model the combined effect of conduction between the air masses separated by the surface, as well as long wave radiation and convection between the surface and the air mass in contact with it [9, 10]. We refer to parameter values so obtained as "ASHRAE values". However, there is uncertainty in ASHRAE values. First, the room capacitance is unknown since air as well as furniture and other objects in the room all contribute to it. Second, information on wall and window construction and material is not always easy to obtain for existing buildings due to poor record keeping. Third, due to cracks in windows, walls, and air gaps around closed doors, the effective resistance of an window or a wall is likely to be lower than what is inferred from construction data. Finally, if a window or a door is open, the effective resistance could be far lower than the resistance estimated for a closed window or door. Therefore, even if R-C network models are adequate, there is a need to *identify/estimate* R-C parameters from measured data; in other words, to *calibrate* the model.

The model calibration problem is not trivial, mostly because of the need to estimate values of R-C parameters as opposed to obtaining a system realization to mimic observed input-output behavior. There are several reasons for estimating these parameters. First, R-C values have intuitive, physical meaning. Second, one can check if the identified parameters values are reasonable—and therefore if the model is—by comparing against ASHRAE values. This provides a sanity check and can help unearth potentially grossly inaccurate system identification. We will see an example of this in Section 3.3. Third, a trend in identified parameter values over time, when identification is carried out repeatedly, can reveal wear and tear of the building fabric or damage in wall insulation. Fourth, since R-C modeling paradigm is prevalent in the HVAC community, it is useful simply as a common language.

Even though powerful state-space identification methods such as the subspace

method [11] can be used to identify the system matrices in an arbitrary state space, they are not useful for estimating R-C parameters. Parameter estimation techniques for transfer functions, such as least-squares ARMA model identification methods [11] and adaptive identification methods [12], are also inadequate. These methods can estimate coefficients in the polynomials that describe the transfer functions. However, obtaining R-C parameters from identified transfer function polynomial coefficients is quite challenging due to the complex relationship between the two. Moreover, the majority of commercial buildings in the U.S. use forced ventilation, in which case the common control signals at a zone are supply air temperature and flow rate. This introduce bilinearity into the model. Although there exists methods to estimate bilinear models (such as [13]), these methods estimate a realization in some arbitrary state space basis and not parameters in the model.

The paper makes three contributions on the model calibration question that are relevant to model based control of indoor climate of a room. First, we answer the minimal model complexity question by a comparison of response between a R-C network model of high state dimension and low-order models of same class. When the mass flow rate of supply air is held constant, the model is LTI. Frequency domain comparisons show that a second-order model with 8 parameters is capable of reproducing the input-output behavior of the full-scale model of the room (with 13 states and 32 parameters) with a high degree of accuracy. The conclusion is that a large number of R-C parameters is not required; a model with two capacitor suffices. This answers the question how many R and C parameters are needed if we accept R-C network models as adequate representation of reality. Existing work on the minimal model complexity question usually starts from modeling the heat transfer through a single surface. Models of multiple surfaces that make up the room are then combined to construct the model of the room. See for example, Mathews et al. [9], Gouda et al. [14, 15], Fraisse et al. [10], and Braun et al. [16]. One drawback of this approach is that the order of the zone becomes large when individual building elements are combined. This problem can be avoided by directly constructing a low-order model for the whole zone, which is the idea adopted in [17, 18] as well as in this paper. Madsen et al. [17] develop a second-order model for a house, and study the time domain performance of the model. However, comparison between the second-order model and possible higher-order models, as well as frequency domain analysis, are not discussed. Bacher et al. [18] provides a zone model selection procedure. First, a set of feasible models of different orders is chosen. A maximum likelihood method is used to determine the unknown parameters in each model and the likelihood function

value corresponding to the optimal parameters is recorded and compared among different model structures. The selection starts from the simplest model, and stops when the likelihood function value does not increase significantly. However, the relation between the likelihood function value and the predictions accuracy is not clear. Their model selection approach is similar to ours in that they start from the simplest model and increase complexity gradually, but we differ in model evaluation.

The second contribution of the paper is that it answers the model calibration question, with a qualified "yes" to the adequacy of R-C network models for modeling a room's thermal dynamics. The R-C parameters are estimated through numerical minimization of a prediction error cost, with data collected from a room in a building in the University of Florida campus. The conclusion of the study is that data collected from forced response tests that ensure certain features can lead to reliable calibration that can reproduce input output behavior reasonably well in a wide range of scenarios. We provide guidelines on what features the data should contain so that it leads to accurate parameter estimation, and what kind of forced response tests are needed to ensure that the data has those features. It turns out that data collected from building during usual operation, even when the inputs are persistently exciting, can lead to grossly "wrong" identification of the parameters. These parameters can nevertheless reproduce input-output behavior under some commonly encountered, but special, conditions in the building. That the parameters are incorrect becomes obvious only when the model is asked to predict outputs in scenarios that are somewhat uncommon at present, namely, when each room's climate differs significantly from their neighboring rooms'. It should be emphasized that such situations will arise if control algorithms are employed to control room-level climate based on occupancy to minimize energy use [2, 19, 4].

In many existing papers about the model calibration problem, such as Madsen et al. [17] and Bacher et al. [18], the heat gain from heating/cooling devices enters the model in a linear fashion. However in this paper, we consider the bilinear scenario, which is common in the U.S. as mentioned before. This renders many tools for linear systems inapplicable. Agbi et al. [20] focus on the identifiability of the model and experiment design. They simulate a 13-zone model with 52 states and 150 parameters using a number of different inputs, which includes step, multi-sinusoidal, and random signals. An information metric is computed to evaluate the difficulty of parameter estimation. Their result shows that some parameters are hard to identify even with inputs that are often deemed rich. This result agrees with the results of our study.

The third contribution is that we establish that one needs a door status sen-

sor to perform model-based control. When the door is open, the convective heat transfer through the open area has a significant effect on the room's thermal dynamics. If one calibrates the room model with door-closed data, the prediction for the door-open case is poor. We show that the effect of the open door can be adequately modeled as an additional resistor between the hallway and the room. Once the door status is known, the control algorithm can switch between the door-closed model and the door-open model, which ensures good prediction at all times. Without a sensor to measure door open/close status, application of MPC to indoor climate control is not likely to be feasible.

A preliminary version of this work was presented in [21]. The focus of this manuscript is on model structures with different boundaries, so that different effects from different surrounding room temperatures can be captured; while the models analyzed in [21] were not able to achieve this task. The model calibration problem is also more challenging since many more parameters are involved in differential boundary models. Only three data sets were used for calibration and validation in [21] and they are inadequate to identify the parameters in the differential boundary models. In this manuscript we use two additional data sets to test prediction accuracy of various model structures. The calibration algorithm presented in [21] was based on direct search, while the one presented here is based on a quasi-Newton search method, and involves additional steps. Compared to [21], this manuscript involves more experiments and detailed explanations.

The rest of the paper is organized as follows: Section 2 introduces the structure of both high- and low-order models and presents a comparison; model calibration is discussed in Section 3, where identifiability, cost function formulation, and calibration method and results are presented; Section 4 shows the effect of an open door on the thermal dynamics of the room and provides a way to model the effect; finally, Section 5 concludes our work and discusses avenues for future work.

2. Model structure (Q1)

In this section, we will describe the structure of several models of a room in the class $\mathscr{M}_{RC}(n,q,p)$, enumerating their states, inputs, and parameters.

Consider a typical room in a multi-room building with VAV system, shown in Figure 1. It is separated from the ambient environment through an external wall and a window, and from five internal spaces (the room above, the room below, one room on each side, and a hallway) through internal walls and a door to the hallway. The external wall and internal wall are two types of walls due to different

construction. The major heat transfer mechanisms include the following: (1) heat exchange through external and internal walls, windows, roof, and ceiling; (2) heat convection with outside air due to the air supplied to and extracted from the room by the HVAC system; (3) solar radiation through the window; (4) heat gain from occupants and equipment, and (5) infiltration and exfiltration.

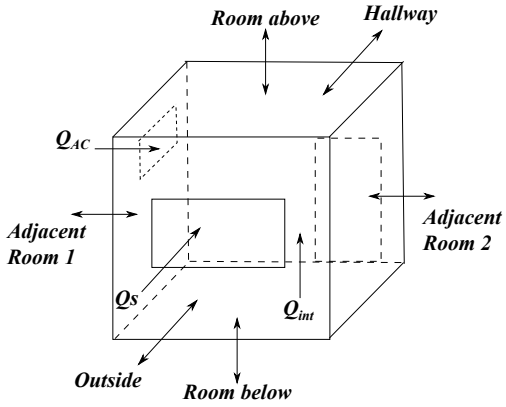

Figure 1: Room configuration studied.

This room configuration is quite common in commercial buildings. A model for such a room also encompasses the model for a room that does not share a wall with the outside, or does not have an window. The only other common case that is not captured by the chosen room configuration is when the room is located on the ground floor so that the temperature on the other side of the floor is the ground temperature.

The following assumptions are made throughout the paper. First, the air inside the room is well mixed, so that we have one uniform temperature in the room. Second, no exfiltration occurs. Exfiltration affects the energy consumption but not the temperature dynamics; it hardly matters if the air leaving the room is leaving through a return air grille or through cracks in an window. Third, infiltration does not occur if the door is closed. Most commercial buildings are maintained at positive pressure to preclude the possibility of infiltration from the outside environment through cracks, so assuming infiltration does not occur is reasonable.

A high-order full-scale model and a second-order model will be described. In Section 2.1 and 2.2, we focus on the door-closed case, where there is only surface between the room and the hallway. When the door is open, there may be

significant heat transfer between the room and the hallway that can't be captured by a R-C network model of the wall separating them. Model of this general case is discussed in Section 2.3. Also, our main interest is on "differential boundary" model structures, in which each adjacent room temperature is a distinct input, so that the different effects of different surrounding room temperatures can be captured. However, a lumped-boundary model will also be briefly introduced and be used as an intermediate step in model calibration (Section 2.2.2). In the sequel, we denote the parameters in the model as θ, the state vector as x, and the input vector as u. We now describe the various models.

2.1. Full-scale model

The full-scale model of room thermal dynamics is constructed by combining R-C models for individual surfaces, a room capacitance, and heat gains. Heat transfer through each surface, except for a window, is modeled as 3R-2C, which is inspired by the results in [14]. Since windows have very low heat capacitance, a window is modeled as a single resistor. Each surface element is then connected to the room "node" to form a R-C network model. An additional capacitor is included to model the heat stored by the air and other objects in room. The heat gains include enthalpy change due to supply and exhaust air and internal heat gains. The structure of the full-scale model is shown in Figure 2. Since each wall

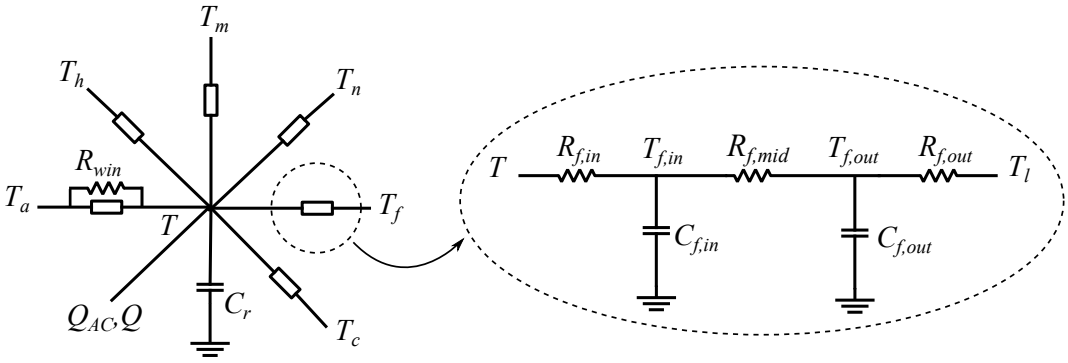

Figure 2: Full-scale model structure for the door-closed case.

is a 3R-2C component, there are two states for each wall. Together with the room temperature T, we have a 13-state vector:

$$x = [T, T_{f,in}, T_{f,out}, T_{c,in}, T_{c,out}, \ldots]^T,$$

where $T_{*,in}$ are the temperatures of the inside nodes of surface $*$, and $T_{*,out}$ are the temperatures of the outside nodes of of surface $*$. Of the two wall nodes, the inside node refers to the node that is closer to the room, while the outside node refers to the one that is closer to the surrounding space. The inputs that affect the room temperature are ambient (outside the building) temperature, temperatures of the surrounding rooms and hallway, heat gain inside the room, and enthalpy exchange due to ventilation. We define the input vector to be

$$u = [T_a, T_f, T_c, T_m, T_n, T_h, T_s, m, Q]^T,$$

where T_a is the temperature of the ambient, T_f is the temperature of the space below the floor, T_c is the temperature of the space above the ceiling, T_m and T_n are the temperature of two adjacent rooms to the side, T_h is the temperature of the hallway, Q is the heat gain from occupants, appliances etc., m and T_s are the flow rate and temperature of supply air, respectively. For resistors and capacitors, we introduce the notation $R_{*,*}$ and $C_{*,*}$. The first subscript indicates which surface it belongs to. The second subscript indicates the position of the resistor/capacitor. We use *in* for the ones next to the room node, *mid* for the ones in the middle, and *out* for the ones next to the outside of the surface. For example, $R_{f,in}$ means the inside resistor of the floor. With this notation, the dynamics of T is:

$$\begin{aligned} C_r\dot{T} =& (-\frac{1}{R_{win}} - \frac{1}{R_{a,in}} - \frac{1}{R_{f,in}} - \frac{1}{R_{c,in}} - \frac{1}{R_{n,in}} - \frac{1}{R_{m,in}} - \frac{1}{R_{h,in}})T_i + \frac{T_a}{R_{win}} + \frac{T_{a,in}}{R_{a,in}} + \\ & \frac{T_{f,in}}{R_{f,in}} + \frac{T_{c,in}}{R_{c,in}} + \frac{T_{n,in}}{R_{n,in}} + \frac{T_{m,in}}{R_{m,in}} + \frac{T_{h,in}}{R_{h,in}} + Q_{AC} + Q, \end{aligned} \tag{1}$$

where Q_{AC} is the net heat gained by the room due to the ventilation air circulating through it. This heat gain equals to the enthalpy of supply air minus the enthalpy of exhaust air [22]:

$$\begin{aligned} Q_{AC} &= h_{in} - h_{out} \\ &= mC_p(T_s - T) + m\left[W_s(h_{we} + C_{pw}T_s) - W(h_{we} + C_{pw}T)\right] \end{aligned} \tag{2}$$

where C_p is the specific heat of air, C_{pw} is the specific heat of water vapor, h_{we} is the evaporation heat of water at 0 °C, and W_s and W are the humidity ratios of supply air and room air. In this paper we assume that $W_s(t) \equiv W(t)$ (see Remark 1). Then Q_{AC} can be simplified to $Q_{AC} = m(T_s - T)(C_p + W_s C_{pw})$. In addition, C_p is

an order of magnitude larger than the $W_s C_{pw}$ terms under normal indoor condition, so we ignore them in the model. As a result, in the remainder of the paper we use

$$Q_{AC} = mC_p(T_s - T). \tag{3}$$

The dynamics of the wall nodes of each wall have similar structure, and are obtained by heat balance:

$$\begin{aligned} C_{*,in}\dot{T}_{*,in} &= (-\frac{1}{R_{*,in}} - \frac{1}{R_{*,mid}})T_{*,in} + \frac{T}{R_{*,in}} + \frac{T_{*,out}}{R_{*,mid}} \\ C_{*,out}\dot{T}_{*,out} &= (-\frac{1}{R_{*,mid}} - \frac{1}{R_{*,out}})T_{*,out} + \frac{T_{*,in}}{R_{*,mid}} + \frac{T_*}{R_{*,out}} \end{aligned} \tag{4}$$

For each surface, there are three resistances and two capacitances, which results in 30 parameters for all six surfaces. Together with the room capacitance and window resistance, there are a total of 32 parameters in the model. Thus, the full-scale model is of the class $\mathscr{M}_{RC}(13,9,32)$.

Remark 1. *The assumption that $W_s(t) \equiv W(t)$ only holds if there is no source or sink for water vapor in the room. We ensure this is the case during model calibration by only collecting data during unoccupied times and ensuring there are no plants or other sources of water vapor. Under such circumstance, the unknown humidity inputs to the model vanish, making parameter estimation easier. Once parameters are calibrated, they can be used for prediction with all inputs presented. This method introduce no loss of generality since inputs do not change the value of parameters.* □

2.2. Second-order models

We now consider low-order models that still employs the R-C network analogy of heat transfer as the full-scale model described above, but with fewer states and parameters. The response of the room temperature T to changes in mass flow rate and temperature of the supply air is usually faster than its response to changes in the surrounding temperatures. To reproduce this two time scales of the process, at least two capacitors are required, which results in a second-order model.

2.2.1. Second-order differential boundary model

To make the model composible, we still assign each surface a different resistor. One capacitor (room capacitance C_r) is used for the thermal mass of the air and

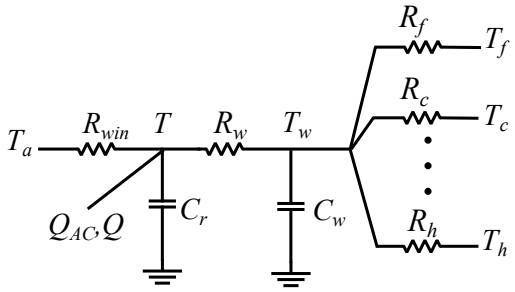

Figure 3: Second-order differential boundary model structure.

other objects in the room, and the other (wall capacitance C_w) is used for the heat capacity of all the walls combined. A different resistors is assigned to each surface. The model structure is shown in Figure 3.

The dynamics of this model are:

$$\begin{aligned} C_r\dot{T} &= \frac{T_a - T}{R_{win}} + \frac{T_w - T}{R_w} + Q_{AC} + Q \\ C_w\dot{T_w} &= \frac{T - T_w}{R_w} + \frac{T_f - T_w}{R_f} + \frac{T_c - T_w}{R_c} + \frac{T_m - T_w}{R_m} + \frac{T_n - T_w}{R_n} + \frac{T_h - T_w}{R_h} \end{aligned} \tag{5}$$

where R_w is assumed to equal to the effective resistance of all the surfaces resistance connected in parallel, i.e., $\frac{1}{R_w} = \frac{1}{R_f} + \frac{1}{R_c} + \frac{1}{R_m} + \frac{1}{R_n} + \frac{1}{R_h}$. In this model, we have 2 states $x = [T, T_w]^T$, the same 9 inputs as in the first-order model, and 1 new parameter C_w. Thus, it is of the class $\mathscr{M}(2,9,8)$.

2.2.2. Second-order lumped boundary model

In this model, all surrounding spaces are combined into one integrated wall and a 2R-1C element is used to model that wall, which leads to the structure in Figure 4. Since the effect of different surrounding spaces are not distinguished in this model, we cannot easily combine individual room models to form a model of the building. Though not composable, this lumped boundary model will help in the model calibration of the differential boundary model, which will be described in Section 3.3.

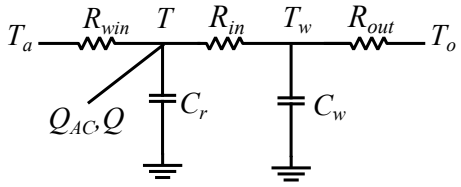

Figure 4: Second-order lumped boundary model structure.

The dynamics of this model are:

$$\begin{aligned} C_r\dot{T} &= \frac{T_a - T}{R_{win}} + \frac{T_w - T}{R_{in}} + Q_{AC} + Q \\ C_w\dot{T_w} &= \frac{T - T_w}{R_{in}} + \frac{T_o - T_w}{R_{out}} \end{aligned} \tag{6}$$

where T_w is the temperature of the wall node. This model has 2 states $x = [T, T_w]^T$, 5 inputs $u = [T_a, T_o, T_s, m, Q]^T$, and 5 parameters $\theta = [C_r, C_w, R_{win}, R_{in}, R_{out}]^T$. Thus, it is of class $\mathscr{M}(2,5,5)$.

2.3. General model for both door-open and door-closed cases

When the door is kept open, a large open area is created, through which large heat exchange between the hallway and the room may occur. We use a resistor to capture this phenomenon. The addition heat transfer term is:

$$Q_h = \frac{1}{R_{od}}(T_h - T) \tag{7}$$

where R_{od} is the effective resistance of open door. The new room thermal dynamics with an open door are given by:

$$\begin{aligned} C_r\dot{T} &= \frac{T_a - T}{R_{win}} + \frac{T_w - T}{R_w} + Q_{AC} + Q + Q_h \\ C_w\dot{T_w} &= \frac{T - T_w}{R_w} + \frac{T_f - T_w}{R_f} + \frac{T_c - T_w}{R_c} + \frac{T_m - T_w}{R_m} + \frac{T_n - T_w}{R_n} + \frac{T_h - T_w}{R_h} \end{aligned} \tag{8}$$

A general model can be obtained by augmenting the door-closed model with the new door-open resistor and a switch, which can be used for prediction in both door-open and door-closed scenarios. To illustrate the idea, the new model structure of the second-order differential boundary model is shown in Figure 5. All

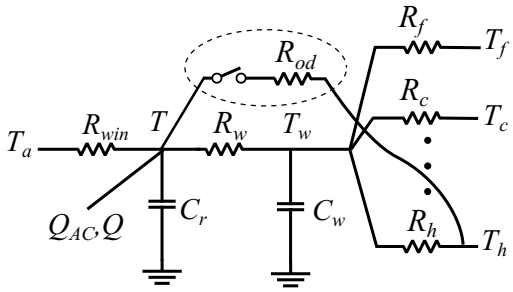

Figure 5: Second-order differential boundary model with a switch and an additional resistor between the hallway and the room to model the effect of open door.

the model structures discussed before can be augmented with resistor and door switch in the same manner to obtain general model. When the switch is closed, the resistor is connected, which models the door-open scenario; when the switch is open, the model changes back to the original door-closed case. Although simple in structure, it turns out that this model is sufficient to provide good prediction; more detail is provided in Section 4.

Remark 2. *Note also that, in all the models mentioned above, the Q_{AC} term is the only nonlinear term in the models. When the supply air flow rate, m, is constant, the Q_{AC} term becomes linear in the state T and input T_s. The system then becomes LTI:*

$$\dot{x} = Fx + Gu, \quad T = Hx, \tag{9}$$

where the state x and input u varies depending on which of the three models is under consideration. The output is the room temperature T since it can be measured, but the output matrix H depends on the choice of model. The matrix F depends on the parameter m. The number of inputs is reduced by one for each class, since m has been moved from being an input to a known constant parameter.

2.4. Frequency domain comparison

In this section, we will compare two models: full-scale model $\mathscr{M}_{RC}(13,9,32)$ and second-order differential boundary model $\mathscr{M}_{RC}(2,9,8)$. We do not deal with calibration in this section. Instead, we first fix the parameters of a full-scale model. Then we construct low order models from this full-scale model through an *ad-hoc model reduction* procedure, whereupon these models are compared among

themselves. The motivation for doing so is the following: if a model obtained from an ad-hoc procedure can reproduce full-scale model behavior, a model of same structure but with parameters calibrated from a more methodological approach will do better.

The parameters in the low-order models are obtained from those in the full-scale model as follows. First, the room capacitance and the resistance of the windows remain the same as in the full-scale model. Second, the integrated wall capacitance is calculated by aggregating each surface as parallel component, i.e., the total capacitance is the sum of capacitances of each surface. Third, since low-order models retain the differential boundary structure, the resistance of each surface is also kept at the same value as in the full-scale model.

To determine the parameters of the full-scale model as well to obtain input signals, a typical office in the second floor of a University of Florida building (Pugh Hall) is used as a prototype. This room, referred to as the *target room* in the sequel, is of the configuration shown in Figure 1. It has a dimension of $15ft \times 15ft \times 9.25ft$, and its climate is controlled by a dedicated VAV terminal box. To determine resistance and capacitance values, we refer to ASHRAE handbook [7] and available construction information.

By examining data from Pugh hall for several months, we find that the supply air temperature has the fastest change rate among all input signals. It can change from its minimum value to maximum value in 5 minutes. Assuming that the largest period we consider is 5 hours, we choose the frequency of interest to be $\frac{1}{5\,hours}$ ($\approx 10^{-4} rad/s$) to $\frac{1}{5\,mins}$ ($\approx 10^{-2} rad/s$). The discrepancy between models was found to increase with increasing value of supply air flow rate. The flow rate is therefore set to $m = 0.1(Kg/s)$, which is the maximum value for the target room.

First, we consider the magnitude frequency response. Among all the inputs, T_s is seen to have the largest gain, and its magnitude frequency response are shown in Figure 6 left. Outside temperature T_a has smaller gain to the output compared to the other inputs, which is consistent with the fact that commercial buildings are usually well-insulated. Gain for surrounding temperatures lie between the gain for T_a and T_s, which is determined by the thermal properties of the surfaces. As shown Figure 6 left, for the range of frequencies deemed of interest, the second order model is almost as accurate as the 13-th order model.

The prediction in response to an input depends not only on the transfer function but also on the magnitude of the input. The variation in T_s and surrounding space temperatures can be significantly different from each other. The error in the

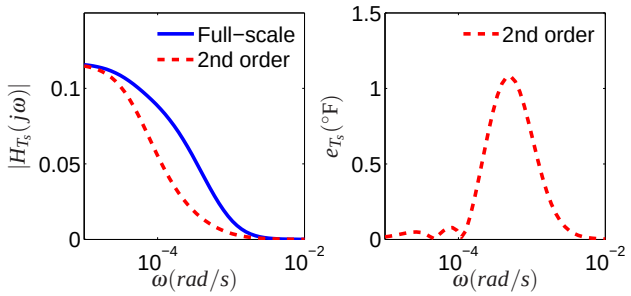

Figure 6: Frequency domain comparison for the input supply air temperature T_s. Left: gain comparison; right: prediction difference.

prediction of T may be different due to the difference in magnitudes of the inputs even though their transfer functions might have the same gain. Therefore we next examine the *prediction difference with respect to i-th input*, which is defined as:

$$e_i(\omega) = \Big| |H_i^{full}(j\omega)| - |H_i^{low}(j\omega)| \Big| \delta u_i, \tag{10}$$

where $H_i^{full}(\cdot)$ and $H_i^{low}(\cdot)$ correspond to the transfer function of full-scale and low-order models from the i-th input to output, and δu_i is the maximum variation in the i-th input. By examining the data from Pugh Hall, we observe that the largest variations in supply air temperature and surrounding room temperatures are 45°F and 10°F. However, if more intelligent control strategy is applied, controller may let the room temperatures vary more to save energy[2, 19, 23, 4]. In that case the temperature difference between adjacent rooms may grow bigger. Thus we use $\delta u_o = 20^\circ$F for the surrounding temperature inputs. The ambient temperature is assumed to have a maximum variation of 30°F in the frequency of interest.

It turns out that the input with most significant prediction difference is still T_s. The difference e_i as a function of frequency for this input is shown in Figure 6 right. We see from the figure that for the second order model, the maximum difference is less than 1.2°F. These conclusions provide us initial confidence in the second-order differential boundary model.

3. Model calibration (Q2)

Since the second-order differential boundary model has far fewer parameters than the full-scale model (8 vs. 32) and the predictive power of the two models is

quite similar, as seen in the previous section, we limit our attention to the second order model $\mathscr{M}_{RC}(2,9,8)$ for the purpose of identification.

3.1. Identifiability

First we address the question of local identifiability of the second-order model $\mathscr{M}_{RC}(2,9,8)$ in the sense defined in [24] for LTI systems. We consider a constant supply air flow rate, which ensures the model is LTI (see Remark 2). Since varying flow rate provides more information, if the model with constant flow rate is identifiable, the model with varying flow rate is likely to be so as well. With constant supply air flow rate, the discrete version of system (5) can be written as:

$$\begin{aligned} x(k+1) &= F(\theta)x(k) + G(\theta)u(k) \\ y(k) &= H(\theta)x(k) \end{aligned} \tag{11}$$

Local identifiability is defined as follows: a model is locally identifiable at θ_0, if in the neighborhood of θ_0 there does not exist two distinct parameters which produce the same input-output behavior. By Proposition 2 in [24], we can determine the identifiability of the model by computing the rank of the information matrix. The precise statement is as follows: consider the matrix

$$S_r(\theta) \triangleq [M(1,\theta), M(2,\theta), \cdots, M(r,\theta)]$$

where $M(j,\theta)$ is the j^{th} Markov parameter of the system (11) organized row-wise:

$$M(j,\theta) = [M_1(j,\theta), M_2(j,\theta), \cdots, M_p(j,\theta)]$$

where $M_n(j,\theta)$ is the n^{th} row of the corresponding Markov parameter. The model is locally identifiable at θ_0 if for large enough r, $rank(\frac{\partial S_r(\theta)}{\partial \theta})$ equals to the number of parameters at θ_0. The minimum number of Markov parameters required r is given by the sum of the observability index and controllability index of the system (11). We let θ_0 take the value discussed in Section 2.4. It turns out that $r = 3$ suffices and the model is locally identifiable at θ_0.

3.2. Cost function formulation

Identification will be performed by minimizing a cost function that captures how well a model (with a given set parameters) predicts measured outputs. To solve the minimization problem applying available numerical optimization tools, we need a discrete time representation of the dynamics. Applying forward Euler

approximation to the continuous-time model (5) and (6), the discretized system model is given by:

$$\begin{aligned} x_{k+1} &= x_k + t_s(Ax_k + Bu_k + f_k) \\ T_k &= Cx_k \end{aligned} \tag{12}$$

where t_s is the sampling time, x_k, u_k is the state and input at $t = kt_s$, $f_k := f(kt_s)$, and T_k is room temperature predicted from the model. For a given model structure with fixed parameters, we define the *prediction error cost J* as:

$$J = \sum_{k=1}^{\tau} (T_k^m - T_k)^2 \tag{13}$$

where T_k^m is the measured room temperature at time k, T_k is the room temperature at time k predicted by the model with a given set of parameter values, and τ is a user-specified time interval. The optimal parameters are those that minimize the prediction error cost J:

$$\hat{\theta}^* = \underset{\theta \in \mathbb{R}^{p+}}{\arg\min}\, J \tag{14}$$

Solution of the minimization problem yields an estimate of the parameters. In this paper we solve the optimization problem (14) numerically using the quasi-Newton BFGS (Broyden - Fletcher - Goldfarb - Shanno) method [25].

3.3. Calibration of second-order differential boundary model

3.3.1. Method

The calibration is carried out in two steps:

1. Identify the room capacitance and window resistance (C_r and R_{win}) by calibrating the second-order lumped boundary model described in Section 2.2.2 by prediction error cost minimization.
2. Identify other 6 parameters in the second-order differential boundary model with C_r and R_{win} values obtained from step 1.

There are two advantages of starting from the lumped boundary model instead of directly calibrating the differential boundary model. First, the room capacitance and window resistance remain the same no matter the surrounding surfaces are combined to one or kept separate. Thus, we can reduce the number of parameters to be estimated in the differential boundary model by using the estimates of those two parameters obtained from calibrating the lumped boundary model. This

Data sets	Date of collection	Description
A	08/02/2011	Normal building operation during summer time.
B	12/05/2011	Normal building operation during winter time.
C	10/24/2011	Forced response test. First the target room was heated up and surrounding rooms are cooled down to generate a temperature difference; then the air supply to the target room is shut down.
D	12/01/2012	Forced response test. The temperature in the surrounding rooms were made to be different from each other.
E	12/03/2012	Forced response test. Besides the features in data set D, the average surrounding room temperatures were made to be different from the target room temperature.

Table 1: Description of experimental data sets (obtained with door closed).

will reduce the dimension of the optimization problem, which lowers the number of possible local minima, as well as computation time. Second, we have fewer parameters to determine in the lumped boundary model, so it is convenient for preliminary investigation.

3.3.2. Data description

For model calibration and validation we collected field data from the target room. All data are collected from 6pm in the evening to 6am next morning, when the room is unoccupied to eliminate the effect of occupancy and humidity dynamics; see Remark 1. The data sets are described in Table 1. The room temperature and a few of the inputs for these data sets are shown in Figure 7.

Richness of the inputs are important for model calibration. Persistency of excitation is a concept that is usually used to describe the richness of input. From [12], an input vector $x(t) \in \mathbb{R}^q$, $q \geq 1$, is called *exciting* over the time interval $[t_0, t_0+\sigma]$, $t_0, \sigma > 0$, if there exists $\alpha > 0$ such that

$$\int_{t_0}^{t_0+\sigma} x(\tau)x(\tau)^T d\tau \geq \alpha I.$$

Among the 5 data sets we collected, inputs in set A, D, and E are *exciting*. However, we'll see in the next section that it is not enough to have exciting inputs to

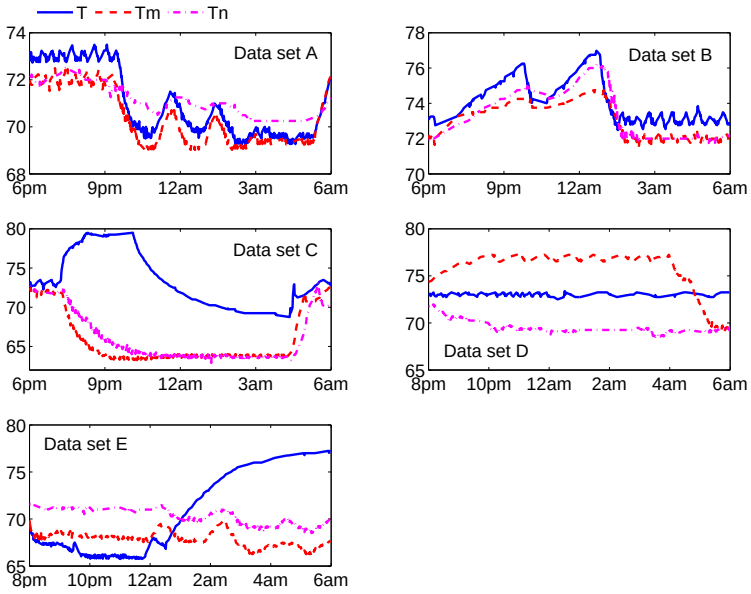

Figure 7: Room temperature T and surrounding spaces T_m and T_n for the 5 data sets. Other inputs including supply air flow rate and other surrounding space temperatures are not shown to avoid clutter.

obtain reliable parameter estimates.

3.3.3. Results

In the initial investigation, data sets A, B, and C are used to calibrate the second-order lumped model. We found that the model with parameters calibrated with data set A predicts the room temperature well for both data set A and B, though set A is collected in summer and set B in winter. However, it fails to predict data set C accurately; see Figure 8. The reason for the failure was found to be the input data; see Figure 7. In the calibration data (data set A), the temperatures of the surrounding spaces are almost the same as that of the target room. So the best fit resistance values are small: the resulting room temperature essentially follows the temperature of the surroundings, leading to small prediction error. In data set B, though the room temperature profile is quite different from that in set A, the surrounding room temperatures are still close to the target room's temperature, so the model with small resistances predicts measurements well again. However, in data set C, the surrounding space temperatures are significantly different from the room temperature, so the model fails to predict measurements well. Only data

set C, obtained from a forced response test, reveals that the calibrated parameter values are incorrect.

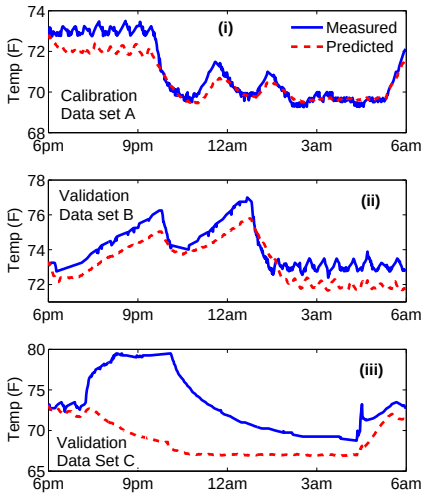

Figure 8: Calibration and (in)validation when data set A is used for calibration; good prediction for data set A and B, poor prediction for data set C.

The best fit parameters for the second-order differential boundary model when data sets C, D, and E are used for calibration, respectively, are shown in Table 2. Also shown in the table for reference are the ASHRAE values for the parameters, which are calculated by the ad-hoc model reduction method described in Section 2.4. Validation simulations are performed with all 5 data sets; results are shown in Figure 9. We found that though the parameter values calibrated from data set C are close to ASHRAE values, they leads to poor prediction in some scenarios; see Figure 9 (v). The reason is that in data set C, surrounding spaces have similar temperatures. A model calibrated to C is not able to differentiate the effects of the surrounding spaces. The inaccuracy in the model becomes evident when it is asked to provide prediction in scenarios where individual surrounding spaces have different temperatures, as in data set E. Data set D does not suffer from this problem since each surrounding space has a different temperature. However, when it is used in calibration, two of the wall resistances, R_f and R_n, turn out to be an order of magnitude lower than their ASHRAE values. The pre-

diction is also poor, as can be seen in Figure 9 (iii, v). We believe the reason for this failure is that even though the surrounding spaces have different temperatures, their average resembles the target room's temperature. Low resistances cause the room temperature to follow the average surrounding temperature, which leads to a small prediction error. As a result, prediction error minimization leads to a small resistance. This is similar to the problem encountered with data set A.

Data set E was created to avoid all of these issues: all surrounding space temperatures are commanded to be lower or higher than the target room's temperature while ensuring that their mean is not equal to the room's temperature; see Figure 7. We see from Figure 9 that the model calibrated with data set E predicts all four data sets (A,B,C,D) well. Also, Table 2 shows that the parameters obtained from data set E are the closest to the ASHRAE values. We can now claim that the model calibrated with data set E is a *low-order differential boundary model with reasonable accuracy*.

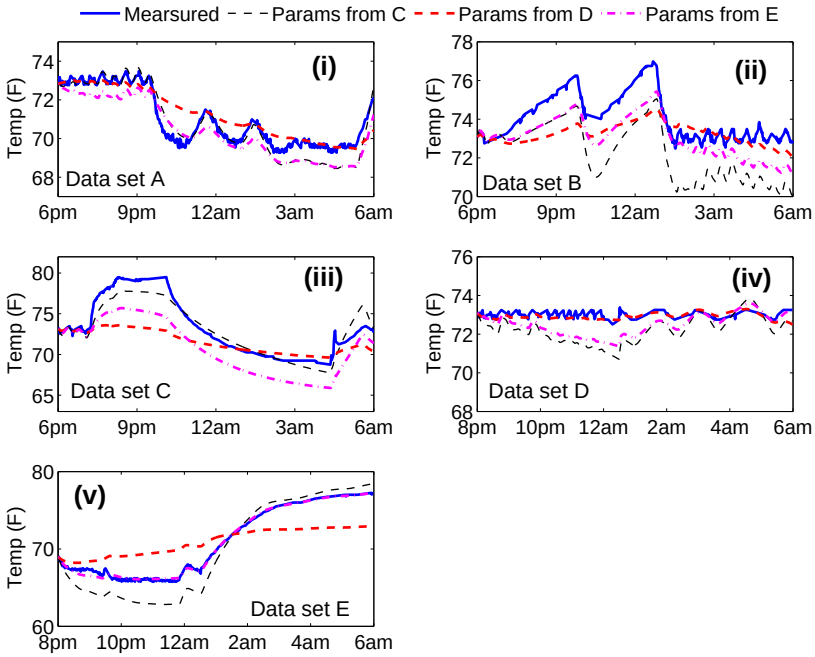

Figure 9: Validation of model calibrated from data set E and invalidation of model calibrated from data set C and D.

Parameter	Data set C	Data set D	Data set E	ASHRAE value
$C_w(J/K)$	1.6×10^6	1.6×10^7	1.1×10^6	$2.2\times10^6-2.3\times10^6$
$R_f(K/W)$	0.39	0.0012	0.78	–
$R_c(K/W)$	0.11	0.42	0.13	–
$R_m(K/W)$	0.2	0.79	0.02	$0.021-0.145$
$R_n(K/W)$	0.025	0.0042	0.02	$0.021-0.145$
$R_h(K/W)$	0.028	0.015	0.0069	-
$C_r(J/K)$	5.2×10^5	1.6×10^6	8.1×10^5	-
$R_{win}(K/W)$	0.059	0.03	0.14	$0.05-0.15$

Table 2: Best fit parameters for the second-order differential boundary model calibrated with different data sets.

4. Effect of open door and its modeling

The first question regarding the open door is whether opening the door can cause significant changes to the room's temperature that the dynamic model needs to take the door into account. We first show that this is indeed the case, by performing another forced response experiment in the target room on Apr. 20th, 2012. During the test, a temperature difference between the room and the hallway is created by commanding the supply air temperature and flow rate in the target room while leaving the corresponding commands for the hallway untouched. The door is kept closed for some time until the temperatures settled down; then it was opened. To see how well our door-closed model (calibrated with data set E) predicts the room temperature in the door open case, the model was simulated with the inputs measured during the experiment. The simulation result is shown in Figure 10. It is clear from the figure that, while the predicted temperature matches the measured temperature well before the door was opened, afterwards it is no longer the case: the predicted temperature decreases while the measured temperature increases. This divergence indicates that the effect of opening the door is significant enough that a model calibrated with "door closed" data is ineffective for predicting "door open" scenarios.

Recall the model structure discussed in Section 2.3. An exhaustive search is used to obtain the best-fit value of R_{od} while all other parameters are fixed at values obtained from calibration discussed in Section 3.3.3. There is only one resistance to determine and we know that it should not be larger than that of a solid surface of

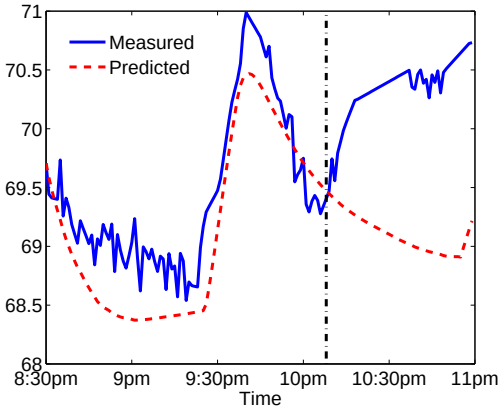

Figure 10: Prediction of door-open data with door-closed model. Vertical line indicates the time door was opened.

the same size, so exhaustive search is both feasible and convenient. The estimated R_{od} turns out to be $0.0043(K/W)$.

To check the door-open model so obtained, we perform a validation door-open experiment in the target room on May 3rd, 2012, in which the door was kept closed first for some time and then opened, but the temperatures of the target room and the hallway were different from those in the previous door-open test (the one used for calibration of R_{od}). The following simulation is performed to compare our model prediction with experimental data: the door-closed model is used for simulation until the door was open; after that we switch to the door-open model. The result is shown in Figure 11. As shown in the figure, we see that the calibrated door-closed and door-open models are able to predict room temperature in an independent test when the room transitions from door closed to door open. It should be noted that even with the calibrated door-closed and -open models, real time prediction of the room's temperature will require knowledge of whether the door is open or closed in any given instant.

5. Summary and future work

We examined two questions regarding models of single room in a commercial building that can be used for predictive control: required model complexity and parameter identification. We examined models of varying complexity within the

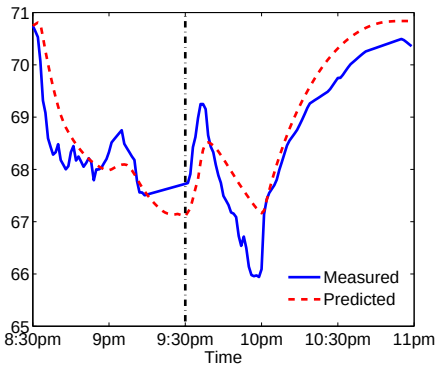

Figure 11: Measured room temperature and temperature predicted from door-open model in validation experiment.

popular class of non-linear R-C network models. We conclude that a second-order model can closely approximate the input-output behavior of the full-scale, 13th order model. Thus, the need for complex models with high state dimension and large number of resistance/capacitance parameters are questionable for control purposes.

The work reported here on parameter identification of low-order models from experimental data has revealed that calibrating the parameters of the R-C network model to normal operation data from a building is likely to lead to grossly inaccurate parameter estimates. A single plot of prediction versus measured data, as often shown in many papers, is virtually meaningless. An incorrect set of parameters may predict certain data sets quite well. Our results identified features that the data should have to enable "correct" identification. These features seem to be possible to ensure only through forced response tests. An algorithm is presented that is able to identify parameter values that lead to a reliable model when such data is used for calibration. By reliable we mean that it is able to reproduce the observed output (room temperature) in a wide range of situations that involve large variations between the output and measurable disturbances. Interestingly, this reliability is seen to come at the cost of accuracy - the prediction error of the temperature is larger for the more reliable model. It raises the question: what is the fundamental limit on the accuracy of a non-linear R-C network model? This is still an open question.

Since specially designed forced response tests are required for calibrating the model of a room, calibrating multiple zones simultaneously could reduce the num-

ber of experiments required. An interesting direction of future work is to design such tests so that the desired data features are guaranteed for each zone.

Our results also reveal that an open door has such a significant effect on room thermal dynamics that one needs to use distinct model structures for the two situations: door-open and door-closed. This has implications on the cost of application of model predictive control in practice, since a sensor to measure the door open/close state will be required to obtain model predictions, and therefore by such a controller.

This work presents model calibration of a room whose all surrounding temperatures can be measured—this is important for simulations since these temperatures appear as inputs to the model. The target room was chosen to be in the second floor of a three-story building to satisfy these requirements. In case of a room in the first floor that is in contact with the building foundation, especially if there is significant ground coupling, the situation is more complicated. It is not clear how to obtain the "ground temperature" measurements required for prediction and identification.

Acknowledgement

The authors would like to thank Physical Plants Division at the University of Florida, and in particular, Peder Winkel, for their help in conducting the forced response tests in Pugh Hall.

References

[1] Energy Use Administration, United States Government, "Electricity explained - use of electricity," 2010. [Online]. Available: http://www.eia.gov/energyexplained/index.cfm?page=electricity_use

[2] P.-D. Moroåÿan, R. Bourdais, D. Dumur, and J. Buisson, "Building temperature regulation using a distributed model predictive control," *Energy and Buildings*, vol. 42, no. 9, pp. 1445–1452, Sep. 2010.

[3] F. Oldewurtel, D. Sturzenegger, and M. Morari, "Importance of occupancy information for building climate control," *Applied Energy*, vol. 101, pp. 521–532, 2012.

[4] S. Goyal, H. Ingley, and P. Barooah, "Occupancy-based zone climate control for energy efficient buildings: Complexity vs. performance," *Applied Energy*, vol. 106, pp. 209–221, June 2013.

[5] American Society of Heating, Refrigerating and Air Conditioning Engineers, "ASHRAE standard 55, thermal environmental conditions for human occupancy," 2010.

[6] R.Kramer, J. Schijndel, and H.Schellen, "Simplified thermal and hygric building models: a literature review," *Frontiers of Archetectural Research*, vol. 1, pp. 318–325, 2012.

[7] American Society of Heating, Refrigerating and Air Conditioning Engineers, "The ASHRAE handbook fundamentals (SI Edition)," 2005.

[8] Carrier corp., "Hourly analysis program (HAP)."

[9] E. Mathews, P. Richards, and C. Lombard, "A first-order thermal model for building design," *Energy and Buildings*, vol. 21, no. 2, pp. 133–145, 1994.

[10] G. Fraisse, C. Viardot, O. Lafabrie, and G. Achard, "Development of a simplified and accurate building model based on electrical analogy," *Energy and Buildings*, vol. 34, pp. 1017–1031, 2002.

[11] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Prentice Hall, 1999.

[12] G. Tao, *Adaptive Control Design and Analysis*. John Wiley & Sons, 2003.

[13] J.-N. Juang, "Continuous-time bilinear system identification," *Nonlinear Dynamics*, vol. 39, no. 1-2, pp. 79–94, 2005.

[14] M. Gouda, S. Danaher, and C. Underwood, "Low-order model for the simulation of a building and its heating system," *Building Services Engineering Research and Technology*, vol. 21, no. 3, pp. 199–208, 2000.

[15] M. Gouda, S.Danaher, and C. Underwood, "Building thermal model reduction using nonlinear constrained optimization," *Building and Environment*, vol. 37, 2002.

[16] J. E. Braun and N. Chaturvedi, "An inverse gray-box model for transient building load prediction," *HVAC&R Research*, vol. 8, no. 1, pp. 73–99, 2002.

[17] H. Madsen and J. Holst, "Estimation of continuous-time models for the heat dynamics of a building," *Energy and Buildings*, vol. 22, pp. 67–79, 1995.

[18] P. Bacher and H. Madsen, "Identifying suitable models for the heat dynamics of buildings," *Energy and Buildings*, vol. 43, no. 7, pp. 1511–1522, Jul. 2011.

[19] Y. Agarwal, B. Balaji, S. Dutta, R. K. Gupta, and T. Weng, "Duty-cycling buildings aggressively: The next frontier in HVAC control," in *Information processing in sensor networks (IPSN)*, 2011, pp. 246–257.

[20] C. Agbi, Z. Song, and B. Krogh, "Parameter identifiability for multi-zone building models," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*. IEEE, 2012, pp. 6951–6956.

[21] Y. Lin, T. Middelkoop, and P. Barooah, "Issues in identification of control-oriented thermal models of zones in multi-zone buildings," in *IEEE Conference on Decision and Control*, December 2012, pp. 6932 – 6937.

[22] S. Goyal and P. Barooah, "A method for model-reduction of nonlinear building thermal dynamics of multi-zone buildings," *Energy and Buildings*, vol. 47, pp. 332–340, April 2012.

[23] S. Goyal, H. Ingley, and P. Barooah, "Zone-level control algorithms based on occupancy information for energy efficient buildings," in *American Control Conference*, June 2012, pp. 3063–3068.

[24] J. V. Doren, P. V. den Hof, J. Jansen, and O. Bosgra, "Determining identifiable parameterizations for large-scale physical models in reservoir engineering," *Proceedings of the 17th IFAC World Congress*, 2008.

[25] C. T. Kelley, *Iterative methods for optimization*. Siam, 1999, vol. 18.