System Identification for Building Thermal Systems under the Presence of Unmeasured Disturbances in Closed Loop Operation: Theoretical Analysis and Application

Donghun Kim[a,*], Jie Cai[b], James E. Braun[a], Kartik B. Ariyur[a]

[a] *School of Mechanical Engineering, Purdue University, West Lafayette, IN, USA*
[b] *The School of Aerospace and Mechanical Engineering, University of Oklahoma, Norman, OK, USA*

Abstract

It is important to have practical methods for constructing and learning a good mathematical model for a building's thermal system in the presence of unmeasured disturbances and using data from closed loop operation. With this goal in mind, this paper presents a mathematical framework that explains the asymptotic behavior of an estimated model under those conditions and that can aid in learning an accurate model. Some analytic results from the literature of system identification are extended and interpreted for building systems. A new identification approach for determining an accurate thermal network (RC) model for a multi-zone building is developed based on the analytic result, and its superior performance over a conventional grey-box modeling approach is demonstrated experimentally.

Keywords: building modeling, grey-box model, system identification, thermal network, disturbances

1. INTRODUCTION

In the past three decades, there has been great interest in applying system identification methods for characterizing building thermal dynamics. Mathematical models obtained through identification processes have been widely used for many purposes. Identified models for buildings can be utilized for shifting cooling loads using the building's thermal capacities, for optimally operating heating ventilation and air conditioning (HVAC) systems using prediction of future loads and for monitoring building energy performance [1; 2; 3; 4; 5; 6; 7; 8; 9; 10].

*Corresponding author
Email address: kim1077@purdue.edu (Donghun Kim)

Considerable research has been performed within the past few decades on methods for obtaining reliable models for building systems. The primary focus of research from the early 1990s has been to select a proper model structure and identification algorithm [11]. Various model structures including black box types, e.g. ARX, ARMAX and state-space forms [12; 13; 14; 15; 16], and grey-box types [4; 17; 18; 19; 20; 21; 22; 23; 24; 25] have been investigated. Many identification algorithms developed from other fields, e.g. signal processing and statistics, have been applied to the field of building science. Popular choices are the least squares method(LS) [13], prediction error methods (PEM) [12], maximum likely-hood methods (MLE) [11] and subspace identification methods (SIM) [12; 26]. Other identification methods, e.g. the identification for long range predictive control [26], extended or unscented Kalman filters [27] and machine learning methods [28; 29] were also applied to building systems. Comparisons between several model structures and algorithms applied to buildings can be found in [26; 30; 31].

Buildings interact with many unmeasured [1]disturbance sources such as occupancy gains, in/ex-filtration and occupant random behaviors, e.g. window/door openings. Furthermore there are many situations where significant disturbance data is missing due to high sensor cost. For instance, information for each zonal plug load which is not negligible for commercial buildings [32, p. 3-5] may not be available, although a whole building plug load can be measured. In addition, it is common to have imperfect information in measurements, e.g. the outdoor air temperature and solar radiation, which also act as unknown disturbances. A typical situation occurs when the building of interest is shaded by its surroundings. In this case, obtained solar data, say from weather services, could be incorrect. It is common to assume schedules for occupancy and equipment in identification processes, e.g. [26] and [18], or to assume that unmeasured disturbances are not dominant, e.g. [33] and [19]. For some buildings equipped with many sensors, the information of dominant disturbances is accessible, and the concern can be alleviated. However, for many buildings in practice, many unknowns exist.

In addition, it is often necessary to obtain a model while maintaining thermal comfort. Therefore, open-loop identification methods, e.g. mode switching of electric heaters with pseudo random binary signals (PRBS), to obtain a better model are limited in practical situations.

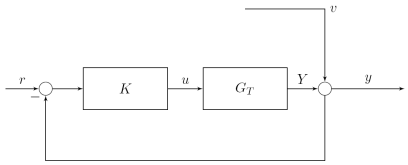
Despite the practical and common issues, there has been little research into the questions of what happens if working data is corrupted by unknown building disturbances and how to obtain a good representation for a building system under the presence of unmeasured disturbances and closed-loop operation. This paper develops a mathematical framework to analyze the behaviors of identified

[1]Uncontrolled external inputs to a system will be called disturbances. We distinguish between *measured* and *unmeasured* disturbances depending on their availability. For example, the ambient temperature is a measured disturbance when the data is available. Otherwise, it is an unmeasured disturbance.

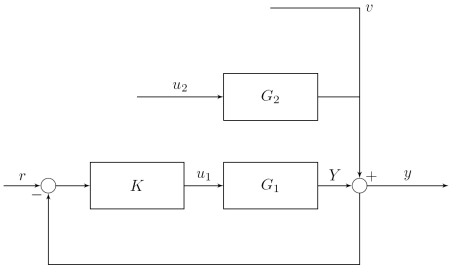
models under those conditions and provides a new identification approach for multi-zone buildings. Sections 2.1 and 2.2 are devoted to describing the system of interest and mathematical frameworks. In Section 2.3, a fundamental result of classical system identification, that describes how an estimated model will be biased when significant unmeasured disturbances exist, will be reviewed. Section 2.4 will extend the result to building applications. The bias expressions derived in this section are used to analyze the relationship between model performance and training data quality. A new identification algorithm is drawn from the analysis and case study results are shown in Section 3.

2. ANALYSIS OF THE INFLUENCE OF UNMEASURED DISTURBANCES ON IDENTIFICATION

2.1. System of Interest



(a) Block diagram of a closed loop system



(b) System of interest

Figure 1: Control Loop Systems

Fig. 1a shows a *standard* block diagram for a closed loop system. G_T is the system to be identified and K represents a feedback controller which responds to a reference signal, denoted as r. u and y are the control input(s) and output(s) in the closed loop system. The signal v is to describe unmeasured *output disturbances*. Y is the output response minus the disturbances, i.e. $Y(k) = y(k) - v(k)$ for each time step k and $Y = G_T \circ u$.

We are particularly interested in a building thermal system in which y is zone air temperature and u is mechanical/electrical heat rate or equipment mode.

Fig. 1a may not be suitable for the purpose of system identification. Most likely, building models have to include some measured disturbances, such as

the ambient temperature or solar irradiation. Therefore, a more suitable form would be Fig. 1b instead of Fig. 1a, where u_1 and u_2 represent the controlled input and measured input disturbances, respectively. Y is the output response purely driven by the control input, where $Y = G_1 \circ u_1$. We are interested in identifying both G_1 and G_2.

2.2. Assumptions and Mathematical Framework for Analysis

This section summarizes some notations and assumptions to describe the theoretical results in Sections 2.3, 2.4 and the Appendix. Terminologies from [34; 35] were adopted.

It is assumed that the true building dynamic system to be identified, denoted as $\mathcal{P}$, is a linear, discrete-time stochastic system having the following form:

$$\mathcal{P} \quad : \quad y(k) = G_1(z)u_1(k) + G_2(z)u_2(k) + v(k) \tag{1}$$

z^{-1} is the backward time shift operator such that $z^{-1}x(k) = x(k-1)$ for any sequence of x. $G_1(z)$ and $G_2(z)$ are stable rational polynomials in z. The assumption of a linear and stable building system is widely adopted, in almost all papers using grey-box based thermal network model structures, despite some nonlinearities associated with convective heat transfer coefficients and radiative heat interactions through windows, as long as u_1 is mechanical/electrical heat rate. v represents the unmeasured (output) disturbances. From Fig. 1b and (1), note that we are not looking at unmeasured heat gains [kW] but looking at their aggregated influence on zone air temperature(s) y [^oC]. This unique view on building disturbances provides significant benefits in analysis and applications: 1) a variety of building heat sources can be replaced with a single process (for a single zone), 2) the output disturbance would be a low pass filtered process through the dynamics of buildings, and 3) together with the assumption described in the next paragraph, the identification problems in practical buildings can be analyzed within a standard mathematical framework of system identification.

Our key assumption for handling the *output* disturbance v is the following.

$$v(k) = H_T(z)e(k) \tag{2}$$

where e is a zero-mean white noise process such that $E(e(k)e(s)) = \sigma^2\delta(k-s)$ for all s and k[2]. H_T is a monic, minimum-phase and stable transfer function. The monic assumption is introduced for normalization reasons and the minimum-phase is to ensure its stable invertibility. Those are standard assumptions in classical system identification books, e.g. [34; 35]. Note that we regard the output disturbance as a filtered process with the filter of H_T, rather than a white noise process. Depending on the shape of H_T in the frequency domain, one can imagine that (2) can generate a variety of disturbances, although it can

[2] δ is the Kronecker delta such that $\delta(k-s) = 1$ only if $k = s$. Otherwise it is zero.

not completely characterize all possible disturbances. Therefore, the color noise assumption together with the output disturbance treatment are anticipated to cover many of building heat sources in practical buildings. If the assumption fails for some cases, the analysis presented in this paper is not valid. However, it is argued that the assumption is versatile enough for most practical purposes [35] in the system identification community.

We restrict parametric models to having the following form.

$$y(k) = G_\theta(z)u(k) + H_\theta(z)\epsilon(k), \tag{3}$$

where ϵ is a zero-mean white noise process and $\theta \in \mathbb{R}^d, d \in \mathbb{N}$. A model structure that maps parameters θ to a model is denoted by $\mathcal{M}$. Examples are ARX, ARMAX and any grey box types.

We denote identification methods, e.g. PEM and SIM, as $\mathcal{I}$ and experimental conditions as $\mathcal{E}$, respectively. $\mathcal{E}$ describes how data is obtained. In this paper, we are interested in data obtained from closed-loop operation.

It is important to mention that in this paper we are interested in system identifiability (SI) rather than parameter identifiablitiy (PI). SI concerns whether an estimated model with $\mathcal{M}$, $\mathcal{I}$ and $\mathcal{E}$ approaches $\mathcal{P}$ as the number of data points goes to infinity [34; 36], while PI concerns whether estimated parameters, e.g. R and C in thermal network models, converge to true parameter values.

2.3. A Brief Review on a Result from a Classical System ID (PEM)

This section reviews a theoretical result of an identification method, PEM, which is relevant for the analysis of the identification with closed loop data. The system of interest is shown in Fig. 1a.

PEM searches parameters, denoted as θ, by minimizing a norm of one-step ahead prediction errors, denoted as ϵ. A standard objective function of PEM, denoted as V, with N data is

$$V(N;\theta) := \frac{1}{N}\sum_{k=1}^{N} tr(\epsilon(k)\epsilon^T(k)), \tag{4}$$

where tr and the superscript T represent the trace and transpose of a matrix, respectively.

Under the assumptions of Section 2.2, the objective function of PEM converges in *probability one* to

$$\frac{1}{N}\sum_{n=0}^{N} tr(\epsilon(k)\epsilon^T(k)) \rightarrow trE(\epsilon(k)\epsilon^T(k)). \tag{5}$$

Ljung [35, pp. 263-265] showed, for a single-input single-output (SISO) system, the objective function can be expressed as

$$\min_\theta \; E(\epsilon^2(k)) = \min_\theta \frac{1}{2\pi}\int_0^{2\pi} \Phi_\epsilon(w)dw, \tag{6}$$

where

$$\begin{aligned}\Phi_\epsilon &= \frac{|G_T - G_\theta + B_\theta|^2}{|H_\theta|^2}\Phi_u + \frac{|H_T - H_\theta|^2(\Phi_e - \frac{|\Phi_{ue}|^2}{\Phi_u})}{|H_\theta|^2} + \Phi_e, \\ B_\theta &= (H_T - H_\theta)\frac{\Phi_{ue}}{\Phi_u}.\end{aligned}$$

where H_θ is an output disturbance model as shown in (3), Φ_i and $\Phi_{i,j}$ are power spectrum of a process i and cross power spectrum of i and j processes, respectively. Frequency dependency w of all terms in the equation are omitted for simplicity in notation.

Consider a simple case where the disturbance model is fixed, say H_0, in order to investigate what happens for closed loop operation. From (6), it is clear that PEM tends to minimize

$$\int_0^{2\pi} |G_T - G_\theta + B_0|^2 \Phi_u dw, \tag{7}$$

where $B_0 = (H_T - H_0)\frac{\Phi_{ue}}{\Phi_u}$. Therefore, the asymptotic estimated model, denoted as $G_{\hat{\theta}}$, will be $G_{\hat{\theta}} = G_T + B_0$ if the range of $\mathcal{M}$ is large enough and $u(k)$ is *persistently exciting* (PE) for any orders, i.e. $\Phi_u(w) > 0, \forall w$. This is clearly not system identifiable. We denote $\tilde{G} := G_T - G_{\hat{\theta}}$ as an asymptotic bias. Then

$$\tilde{G} = -(H_T - H_0)\frac{\Phi_{ue}}{\Phi_u}. \tag{8}$$

Note that (8) relates quality of an identified model to quality of data and will provide an analytic tool to judge the data quality and methods to improve the model for the system of Fig. 1a.

2.4. Analysis of the Bias of a Model for a Building System Caused by Unmeasured Disturbances

We will generalize the expression of the asymptotic bias (8) for SISO systems to MIMO systems, interpret it for buildings, and then use it for designing a new identification approach in Section 3. Consider the system shown in Fig. 1b where the working data for model estimation consists of both control inputs u_1, and measured disturbances u_2. As mentioned, for building modeling, we are also interested in questions of how the model associated with measured disturbances, i.e. the model for G_2 in Fig. 1b, is affected by v and how to improve the model for both G_1 and G_2. In order to answer those questions, a generalized version of (6) for the system was derived in the Appendix.

For a case when $y(k), u_1(k), u_2(k) \in \mathbb{R}$, the asymptotic objective function of a PEM is

$$\min_\theta \; E(\epsilon^2(k)) \quad = \quad \min_\theta \frac{1}{2\pi}\int_0^{2\pi} \Phi_\epsilon(w) dw \tag{9}$$

where

$$\begin{aligned}\Phi_\epsilon &= |\tilde{G_2} + \tilde{G_1}\Phi_{12}\Phi_2^{-1} + \tilde{H}\Phi_{e2}\Phi_2^{-1}|^2\frac{\Phi_2}{|H_\theta|^2} \\ &+ |\tilde{G_1} + B_1|^2\frac{(\Phi_1 - \Phi_{12}\Phi_2^{-1}\Phi_{21})}{|H_\theta|^2} \\ &+ |\tilde{H}|^2\frac{\Phi'_e}{|H_\theta|^2} + \Phi_e.\end{aligned}$$

where $\tilde{G_1}, \tilde{G_2}$ are the model bias, and $\tilde{H}$ is the model mismatch between the true disturbance dynamics and modeled disturbance dynamics. In other words, $\tilde{G_1} = G_1 - G_{1,\theta}$, $\tilde{G_2} = G_2 - G_{2,\theta}$ and $\tilde{H} = H_T - H_\theta$. See (42) in the Appendix for its derivation and other notations. $\tilde{G_1}, \tilde{G_2}, \tilde{H}$ and H_θ are dependent on θ, and all spectrums are determined by estimation data.

For analytical simplicity, focus on a simple case where the disturbance model is fixed. From (9), it is clear that PEM tends to minimize

$$\int_0^{2\pi} |\tilde{G_2} + \tilde{G_1}\Phi_{12}\Phi_2^{-1} + \tilde{H}\Phi_{e2}\Phi_2^{-1}|^2\Phi_2 + |\tilde{G_1} + B_1|^2(\Phi_1 - \Phi_{12}\Phi_2^{-1}\Phi_{21})dw. \quad (10)$$

Clearly 0 is the minimum of (10) and under mild conditions, a minimizer satisfies the following.

$$\tilde{G_1} = -B_1 = -\tilde{H}(\underbrace{\Phi_{e1} - \Phi_{e2}\Phi_2^{-1}\Phi_{21}}\underbrace{)(\Phi_1 - \Phi_{12}\Phi_2^{-1}\Phi_{21}}\,)^{-1} \quad (11)$$

$$\tilde{G_2} = -\tilde{G_1}\Phi_{12}\Phi_2^{-1} - \tilde{H}\Phi_{e2}\Phi_2^{-1} \quad (12)$$

(11) and (12) are the concluding equations that express the asymptotic bias for estimated models of G_1 and G_2. Their interpretations are as follows.

Define a process, denoted by $u_{1|2}$, to be

$$u_{1|2}(k) := u_1(k) - \Phi_{12}(z)\Phi_2^{-1}(z)u_2(k). \quad (13)$$

Then it is easy to see

$$\begin{aligned}\Phi_{1|2,2} &= 0 \quad (14)\\ \Phi_{1|2} &= \Phi_1 - \Phi_{12}\Phi_2^{-1}\Phi_{21}.\end{aligned}$$

This implies the process $u_{1|2}$ is the uncorrelated component of u_1 with respect to u_2. Therefore the second bracketed term of (11) is the power of the signal. $u_{1|2}$ may be thought of as a portion of the control effort to reject the unmeasured disturbances rather than u_2. Note that when u_1 and u_2 are significantly correlated, e.g. $u_1 \approx u_2$ as an extreme case, the second bracketed term is close to 0. Furthermore the first bracketed term is the cross spectrum between e and

$u_{1|2}$, i.e. $\Phi_{e,1|2}$. This can be seen by taking the z-Transform of the correlation function between e and $u_{1|2}$.

With these in mind, (11) means that the asymptotic inconsistency for G_1 appears due to

- the mis-specified disturbance model
- the correlation between unmeasured disturbances and the component of u_1 that is not correlated to measured disturbances
- the correlation between u_1 and u_2.

$\tilde{G}_2$ of (12) represents the asymptotic bias for G_2. It can be clearly seen by the formula, unlike SISO cases, that the model mismatch for G_1, i.e. $\tilde{G}_1$, directly affects the bias for G_2. In other words, inconsistent estimation of G_1 deteriorates the model quality of G_2. Therefore one can not expect a good model for G_2 if the model for G_1 is bad, unless the correlation between the two inputs is weak or the power of u_2 is high enough for all frequencies. Conversely one can improve the model quality for G_2 by improving the model for G_1. In summary, the asymptotic inconsistency for G_2 appears due to

- the mis-specified disturbance model
- the correlation between the measured disturbances and the unmeasured disturbances
- the bias of G_1.

So far, we have shown that the correlations between u_1, u_2 and v deteriorate the quality of the models. However it is not difficult to show that the correlations between individual components of u_1 when the dimension of u_1 is greater than 1, e.g. for a building served by multiple rooftop units (RTU), also deteriorate the quality of the models. Likewise, the correlation between u_2 components influences the estimated model.

Despite the complex mathematical development, the conclusion to judge or design data for the estimation of building models under the presence of unmeasured disturbances and closed loop operation is extremely simple: "*Try to decorrelate all inputs*". This simple goal provides insights for assessing model quality by looking at purely the training data set or for selecting a better data set for practical situations in buildings. Some examples are listed below.

1. Suppose data for ambient temperature and solar are available. Correlation between the two disturbances can cause inconsistency. Therefore, it is wise to select data where the correlation is low, e.g. cloudy days.
2. Suppose ambient temperature and plug-load data are available. If weather and plug-load data are strongly correlated, the models associated with the inputs are most likely not reliable.

3. Consider an open space building served by multiple rooftop units (RTUs). In this multi-input multi-output system, G_1 is not accurate unless all control inputs are uncorrelated. This makes sense because one can not distinguish which RTU affects which thermostat if they are turned on and off at the same time.
4. Consider the case that data is collected from closed loop operation with a fixed zone air temperature setpoint and the ambient temperature and mechanical heat removal rate are measured as inputs. One can easily conclude that G_1 is not reliable, because the cooling load profile must be correlated to the ambient temperature since the controller would reject the signal of the ambient temperature, unless the building is not significantly affected by the temperature.
5. In the above case, one can not also expect a good model for G_2 because $\tilde{G}_1$ affects $\tilde{G}_2$.

The arguments themselves in the examples are not surprising. However the interesting point is that they can be explained within the mathematical expressions of (11) and (12) which support the validity and physical meaningfulness of the mathematical framework.

Unfortunately, the final equations do not provide a specific metric for determining model quality in real-world applications. This is because, 1) their evaluations require the knowledge of unknown dynamics of disturbances, i.e. $H_T(z)$, and the underlying white noise process of $e(k)$. Nevertheless, they provide qualitative methods and theoretical reasons for assessing model quality by looking at purely the training data set, for selecting or designing a better data set as shown in the examples, and for designing an identification algorithm (see Section 3.3 where they are used to identify and resolve the issue of lack of disturbance model associated with typical grey-box thermal network identification approaches).

2.5. Validity range of the analytic results

Note that we did not assume any feedback control so far, thus the conclusions made in Section 2.4 hold for *any* controller, including with several mode changes and control input saturation.

PEM is a general identification method. When PEM is applied to an ARX, it is the same as LS. Furthermore it is well-known that PEM is asymptotically equivalent to MLE [37] when the noise process, e, is Gaussian. More importantly, most grey-box-based identification approaches also belong to PEM as long as the approaches tend to minimize a norm of simulation errors. Shook, Mohtadi, and Shah [38] also showed that an identification method which minimizes a norm of several step prediction errors, so called *identification for a long-range predictive control*, is asymptotically the same as PEM except for one filter term of its asymptotic description (see [38] and [39]). Therefore the following asymptotic results of PEM can be applied for many identification methods which are relevant to the field of building modeling.

3. APPLICATION: AN IDENTIFICATION APPROACH TO ALLEVIATE EFFECTS OF UNMEASURED HEAT GAINS FOR MIMO BUILDING THERMAL SYSTEMS

In this section, we present an identification approach aiming at improving the accuracy of a grey-box model for a multi-zone building despite the presence of unmeasured disturbances. The analytic result, (11) and (12), in the preceding section is used to identify the problem associated with typical RC network identification approaches and overcome the problem. The system of interest is described in Section 3.1. Section 3.2 elaborates an RC network for the multi-input multi-output (MIMO) building. The proposed approach is summarized in Section 3.3. Section 3.4 describes an experimental design approach aiming at obtaining a better model. Experimental case study results are shown in Section 3.5.

3.1. System Description

An open space area served by multiple packaged units is of particular interest. A rooftop unit (RTU) is a packaged air conditioning unit consisting of a vapor compression cycle, supply air blower, air mixing box and optional economizer and heating element, i.e. an electric heater or gas burner. In general, a thermostat is dedicated to a RTU and turns one or more compressor stages on and off to maintain a local zone air temperature near a setpoint. The supply fan is typically on continuously during the occupied period or can cycle with the compressor during unoccupied times. Most RTUs in the field do not have variable speed compressors or fans controlled by variable frequency drives. Conventional thermostat logic for a RTU employs a relay switch where a compressor stage is switched on when the thermostat temperature reaches a temperature that is the setpoint plus a deadband, and remains on until the temperature is cooled to the sepoint.

To formulate the identification problem, let $n \in \mathbb{N}$ be the number of zones or equivalently the number of RTUs or thermostats. The measured outputs are the thermostat temperatures, denoted as $y \in \mathbb{R}^n$. The manipulated variables are the RTU or compressor stages, denoted as $u \in \mathbb{Z}^n$, and we assume the information of the outdoor air temperature, $T_o \in \mathbb{R}$, is available. Let $[u^T, T_o]^T$ be the measured inputs having the size of m where $m = n + 1$. The dynamics of the system, namely $G_u : u \mapsto y$ and $G_{T_o} : T_o \mapsto y$, in nature is very complex, because it involves a refrigerant cycle, the heat exchange between ventilation air and the refrigerant, inter-zonal air flows that could vary depending on the combination of the supply fan modes and air duct systems, and the time constants of thermostats.

The systems to be identified are G_u and G_{T_o}. It is assumed that measurements are collected under closed-loop operation. In other words, u is determined by a control loop. The most significant challenge to modeling is the lack of available measurement points associated with this application. The IO data for identification contains unknown but possibly significant disturbances (e.g.,

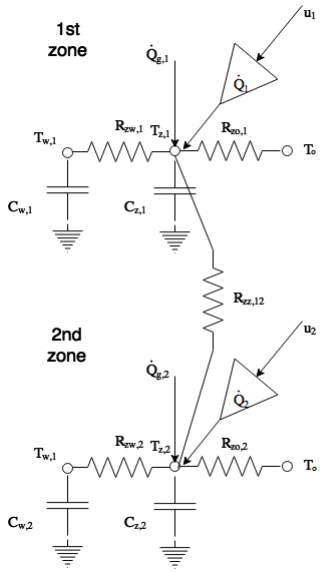


Figure 2: RC network structure for a two-zone building

solar radiation, internal gains), and hence the control inputs are correlated to the unmeasured disturbances due to the thermostat feedback controllers.

3.2. Choice of RC Network Model Structure

We assume a simple model structure in an attempt to capture major dynamics of the complex system. An example RC network for a two-zone (i.e., two thermostat) system is depicted in Fig. 2. Each zone is composed of two thermal capacitances and two thermal resistances, and is served by its own unit having capacity of $\dot{Q}_j$. Zone air nodes are connected by a resistance denoted as R_{zz} in the figure[3].

Corresponding differential equations for the j^{th} zone are;

$$\begin{aligned} C_{z,j}\dot{T}_{z,j} &= \frac{1}{R_{zw,j}}(T_{w,i} - T_{z,j}) + \sum_{k=1,k\neq j}^{n} \frac{1}{R_{zz,kj}}(T_{z,k} - T_{z,j}) \\ &\quad + \frac{1}{R_{zo,j}}(T_o - T_{z,j}) + \dot{Q}_{g,j} + \dot{Q}_j u_j \\ C_{w,j}\dot{T}_{w,j} &= \frac{1}{R_{zw,j}}(T_{z,j} - T_{w,i}). \end{aligned} \tag{15}$$

[3]In reality, the inter-zonal resistance could be time varying depending on the status of supply air fans. Furthermore, the cooling/heating capacity is also time varying depending on the outdoor air temperature, return air wet-bulb temperature and supply air/condenser fan speed. Our intention is to get representative values for those parameters based on measurements.

$T_{z,j}$ is the local zone air (thermostat) temperature and $T_{w,j}$ is a local wall temperature representing averaged behaviors of the building enclosure. $T_o, \dot{Q}_j$ are the outdoor air temperature, and mechanical heat rate for the j^{th} zone. $u_j \in \mathbb{Z}$ represents a RTU stage. $C_{z,j}$ and $C_{w,j}$ represent thermal capacitances of the zone and wall. $R_{zw,j}$ and $R_{za,j}$ are thermal resistances between the zone air and wall, and between the zone air and ambient air, respectively. $R_{zz,kj} = R_{zz,jk}$ is the effective thermal resistances between the k^{th} and j^{th} zone. The total number of interzonal resistances is the number of 2-combinations from n, and hence is $n!/((n-2)!2!)$. $\dot{Q}_{g,j}$ represents unmeasured input disturbances, e.g. occupancy gains, lighting loads and transmitted solar radiation.

The parameters to be estimated are denoted as θ which consists of $C_{w,j}, C_{z,j}$, $R_{zw,j}, R_{zo,j}, \dot{Q}_j$ and $R_{zz,kj}$, $\forall j,k \in \{1,\cdots,n\}$. The dimension of θ is $5n + n!/((n-2)!2!)$.

Based on (15), an n zone system can be expressed as follows.

$$\begin{bmatrix} \tilde{C}_w & \\ & \tilde{C}_z \end{bmatrix} \begin{bmatrix} \dot{T}_w \\ \dot{T}_z \end{bmatrix} = \begin{bmatrix} \tilde{H}_{ww} & \tilde{H}_{wz} \\ \tilde{H}_{zw} & \tilde{H}_{zz} \end{bmatrix} \begin{bmatrix} T_w \\ T_z \end{bmatrix} + \begin{bmatrix} \mathbf{0} & 0 \\ \tilde{\dot{Q}} & \tilde{H}_{zo} \end{bmatrix} \begin{bmatrix} u \\ T_o \end{bmatrix} + \begin{bmatrix} 0 \\ \dot{Q}_g \end{bmatrix} \tag{16}$$

where $\tilde{C}_w = diag(C_{w,1},\cdots,C_{w,n})$, $\tilde{C}_z = diag(C_{z,1},\cdots,C_{z,n})$, $\tilde{H}_{ww} = -diag(1/R_{zw,1},\cdots,1/R_{zw,n})$, $\tilde{H}_{wz} = \tilde{H}_{zw} = -\tilde{H}_{ww}$, $\tilde{\dot{Q}} = diag(\dot{Q}_1,\cdots,\dot{Q}_n)$ and $\tilde{H}_{zo} = [1/R_{zo,1},\cdots,1/R_{zo,n}]^T$. The (k,j) entry of $\tilde{H}_{zz}$ is $1/R_{zz,kj}$ for $k \neq j$. When $k = j$, the j^{th} diagonal entry of $\tilde{H}_{zz}$ is $-(1/R_{zw,j} + 1/R_{zo,j} + \sum_{k=1,k\neq j}^{n} 1/R_{zz,kj})$. Here, $diag(a,b)$ means a diagonal matrix having (a,b) for the entities.

After discretizing the differential equations, one can come up with the following discrete transfer function.

$$T_z(k) = G_u(z)u(k) + G_{T_o}(z)T_o(k) + G_g(z)\dot{Q}_g(k). \tag{17}$$

3.3. Lumped Output Disturbance Identification Algorithm

Our goal is to estimate G_u and $G_{T,o}$ in (17) as accurately as possible based on measurements of u, T_o and T_z without any information of $\dot{Q}_g$.

In this paper, we will call a grey-box based identification approach conventional if its model structure relies on an RC network structure and there are no parameters to estimate dynamics of unmeasured disturbances. Typically conventional approaches seek to find a certain θ by minimizing the sum of squares of simulation errors and have the following model structure.

$$T_z(k) = G_u(z;\theta)u(k) + G_{T_o}(z;\theta)T_o(k) + \epsilon(k). \tag{18}$$

Comparing this to (1) indicate that the true dynamics of unmeasured disturbances, i.e. $H_T(z)$, is identified as the identity matrix I_n in the conventional

approaches. Based on the results in Section 2.4, it is clear that they result in poor models due to the mismatched disturbance model.

More precisely, using (11) and (12), asymptotic bias for G_u and G_{T_o} for the conventional grey-box approaches, are

$$\begin{aligned}\tilde{G}_u &= -\underline{(H_T - I_n)}(\Phi_{eu} - \Phi_{eT_o}\Phi_{T_o}^{-1}\Phi_{T_o,u}) \\ &\quad \times(\Phi_u - \Phi_{u,T_o}\Phi_{T_o}^{-1}\Phi_{T_o,u})^{-1} \end{aligned} \tag{19}$$

$$\tilde{G}_{T_o} = -\tilde{G_u}\Phi_{uT_o}\Phi_{T_o}^{-1} - \underline{(H_T - I_n)}\Phi_{eT_o}\Phi_{T_o}^{-1} \tag{20}$$

where Φ_u, Φ_{T_o} and Φ_e represent the power spectrums of u, T_o and a noise process e. $\Phi_{i,j}$ represents a cross spectrum between i and j processes.

(19) and (20) indicates that the conventional approaches result in poor models, i.e. increased $\tilde{G}_u, \tilde{G}_{T_o}$, due to the discrepancies of disturbance models (underlined terms). This failure is intuitively clear, since the conventional approaches tend to explain input/output relationships without information of $\dot{Q}_g$.

Equations of (11) and (12) indicate that the *only* solution to reduce $\tilde{G}_u$ and $\tilde{G}_{T,o}$ is to lower $\tilde{H}$ in the identification algorithm phase because the other terms are associated with data. The solution approach that is not only estimating physical parameters of θ but also estimating underlying dynamics of the unknown disturbances is termed the lumped (output) disturbance modeling approach (LD) and was presented in our companion paper [40]. The idea of modeling disturbances can be found in most standard system identification algorithms, e.g. PEM on ARMAX model structures and subspace system identification methods, but the LD works on grey-box models with parameter constraints rather than black box models and with *decoupled* output disturbance models as described below.

Since the LD introduced in [40] is limited to a single output system, it can not be directly applied to our systems. Therefore, in this paper, the algorithm is revised to cover the MIMO identification problem by modifying a disturbance model structure. It starts from characterizing disturbances in buildings. Dynamics of a building are influenced by many disturbance sources. However, in the mathematical frame of Section 2.2, the various heat gains were lumped into $v(k)$, and hence one needs to characterize only one signal, i.e. $G_g(z)\dot{Q}_{g,j}$, for each zone. In addition, heat gains may have high power spectrum in a high frequency range, which is difficult to model. The key observation to handle this issue is that if the signal of $G_g(z)\dot{Q}_g(k)$ is modeled, the characterization problem becomes much easier.

More precisely, let

$$v(k) := G_g(z)\dot{Q}_g(k). \tag{21}$$

We call v the lumped (output) disturbances. From (21), v is a filtered signal of unknown heat gains due to building dynamics, i.e. $G_g(z)$. Since building

dynamics are essentially a low pass filter, v has high power spectrum within a low frequency range. A simple low pass filtered signal may be characterized by a pole and DC-gain. Therefore, we model H_T having the following structure.

$$H(z;\rho) = diag(\frac{\rho_{2,1}z^{-1}}{1+\rho_{1,1}z^{-1}} + 1, \cdots, \frac{\rho_{2,n}z^{-1}}{1+\rho_{1,n}z^{-1}} + 1) \tag{22}$$

where $\rho_{1,j}, \rho_{2,j} \in \mathbb{R}$ having bounds of

$$0 < \rho_{1,j} < 1, \tag{23}$$

$$-1 < \rho_{2,j} < 1 \quad , \tag{24}$$

for all $j \in \{1, \cdots, n\}$. The first parameter, $\rho_{1,j}$, and the corresponding constraint make H asymptotically stable and a low pass filter. The second parameter, $\rho_{2,j}$, is to adjust the DC-gain of a low pass filter.

Considering (17) and (22), our final model structure has the following state space form.

$$\begin{aligned} \hat{T}(k+1) &= A(\theta)\hat{T}(k) + B_u(\theta)u(k) + B_{T,o}(\theta)T_o(k) \\ T_z(k) &= [0_n, I_n]\hat{T}(k) + \hat{v}(k) \\ \hat{\zeta}(k+1) &= \mathcal{F}(\rho)\hat{\zeta}(k) + \mathcal{G}(\rho)\epsilon(k) \\ \hat{v}(k) &= \hat{\zeta}(k) + \epsilon(k) \end{aligned} \tag{25}$$

where ρ represents parameters for the disturbance model and

$$\begin{aligned} \mathcal{F}(\rho) &= -diag(\rho_{1,1}, \cdots, \rho_{1,n}), \\ \mathcal{G}(\rho) &= diag(\rho_{2,1}, \cdots, \rho_{2,n}). \end{aligned} \tag{26}$$

$A(\theta), B_u(\theta)$ and $B_{T,o}(\theta)$ are defined by (16) and a discretization scheme for the continuous time dynamic description, e.g. the zero-order hold. $\hat{T}$ is estimated states for zone air and wall temperatures with candidate parameters of θ, $\hat{v}$ is an estimation of true output disturbances with ρ, and $\hat{\zeta}$ is the internal state appearing in converting the transfer function description of the output disturbances (22) to the state space description (25).

A remark on the final model structure (25) compared with other disturbance modeling approaches is that we have additional state $\hat{\zeta}$ augmented to physical state $\hat{T}$, and the dynamics associated with them are completely decoupled. This is to provide flexibility in our disturbance model $H(\rho)$ to fit H_T with a smaller number of parameters. For example, ARX, ARMAX or SIM structures have constraints such that dynamics of a system and disturbance have shared poles while our disturbance model structure does not need to. In this sense, a black box version of (25) is the Box-Jenkins model structure.

Since ϵ in (25) is the innovation process, a simulation error minimization scheme can not be utilized. Therefore, the PEM approach is used for estimating both parameters of θ and ρ based on the state space form of (25). For detailed descriptions to get θ and ρ, see [40].

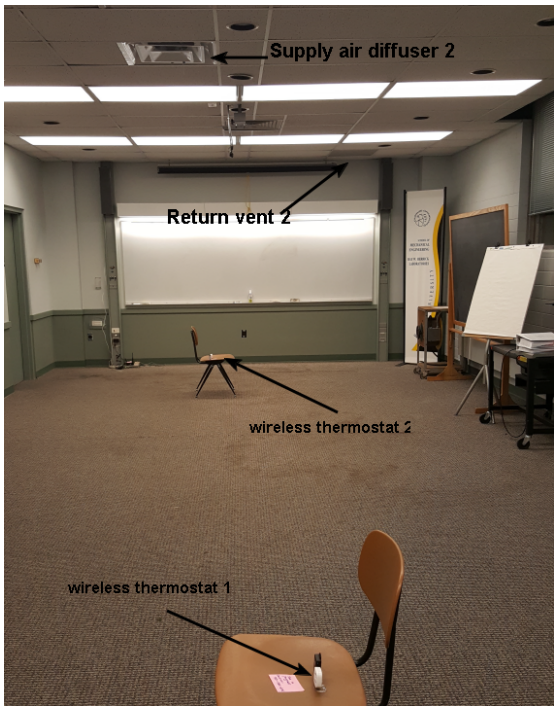


Figure 3: Case study building (supply air diffuser 1 and return vent 1 are not shown)

3.4. Design of Experiment

As mentioned in Section 2.4, it is wise to decorrelate all inputs to reduce unwanted influences of unknown heat gains to an estimated model. From (20), recall that the model mismatch for G_u, i.e. $\tilde{G}_u$, directly affects the bias for G_{T_o} through the correlation between u and T_o. In other words, poor estimation of G_u implies a poor model for G_{T_o}, and vice versa. Likewise, the correlations between control inputs u also deteriorate the quality of the models. It makes sense that one cannot distinguish which RTU affects which thermostat and how much a RTU influences a thermostat temperature when RTUs turn on and off at a same time, i.e. RTU stages are correlated. A standard approach which perturbs setpoints of thermostats randomly within comfort bounds is suitable for this purpose.

3.5. Case Study

3.5.1. Building Description and Experimental Setup

To test the overall identification approach a cooling system for a conference room (about 15 m long, 7 m wide and 3.5 m high) in the Ray W. Herrick Laboratories at Purdue University, IN, U.S. (see Fig. 3) was retrofitted. Two packaged air conditioners (termed RTU1 and RTU2) having different cooling capacities

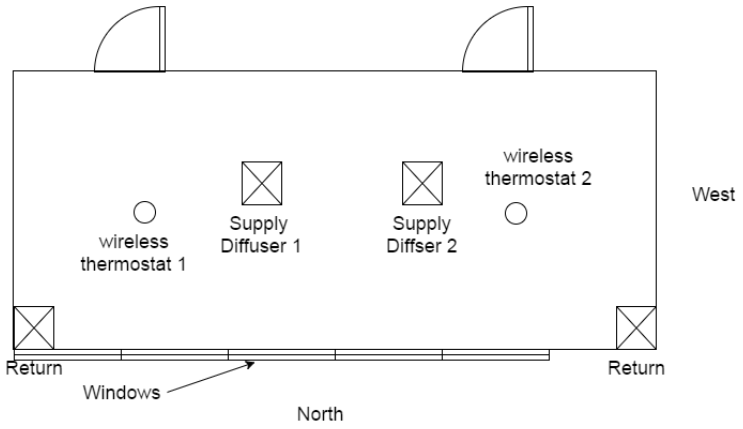


Figure 4: Floor Plan with locations of thermostats, supply and return vents (Purdue Herrick Laboratory)

were installed and the air duct system was reconfigured accordingly. RTU1 is a 1-ton single stage unit with 9 EER and RTU2 is a 2-ton single stage unit with 10 EER. Thermostats, supply and return vents associated with the two units are shown in Fig. 3 and 4. In this system, we have 2 thermostat temperature outputs and 3 measured inputs that are the RTU stages and outdoor air temperature which was obtained from a NOAA (National Oceanic and Atmospheric Administration) website. Note that there is a strong coupling between the two sub-zones (or thermostats), and hence the operation of one unit can influence both thermostat temperatures. Unmeasured heat sources are lighting gains (around 1.5 kW), loads from electric appliances (a small freezer and one laptop computer), in/exfiltration and solar gains though windows. The lights were turned on and off by occupant random behavior.

3.5.2. Results

The thermostat temperatures and each RTU ON/OFF mode, more precisely, the stage run time, were recorded with a five-minute sampling time. The set-points during the daytime were $72 \pm 2^o F$ and $72 \pm 4^o F$ for night time. The input and output data for model training are shown in Fig. 5 for a one-week period. The NOAA outdoor air temperature data was available at a one-hour sampling time, which was interpolated to coincide with the 5-minute sampling time for the other data.

grey-box models were developed using the proposed identification algorithm (LD) and a conventional simulation error minimization algorithm (denoted as Conv) with the same estimation data and the same RC network model described in the preceding section. For detailed descriptions for the conventional approach, see Section 3.3.

Recall that we are interested in the question of system identifiability: "Does the model describe the true system"?. It is important to mention that typical

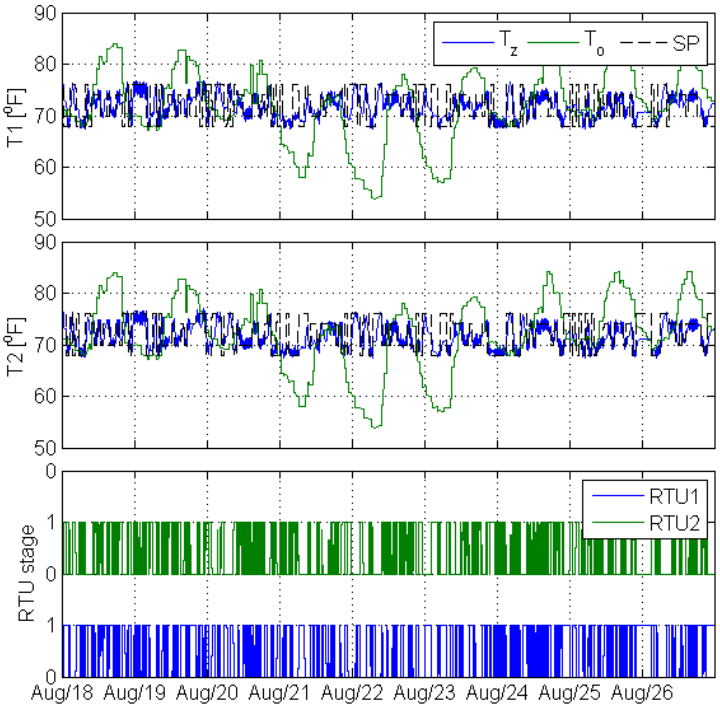


Figure 5: Esimation data set (Purdue Herrick Laboratory)

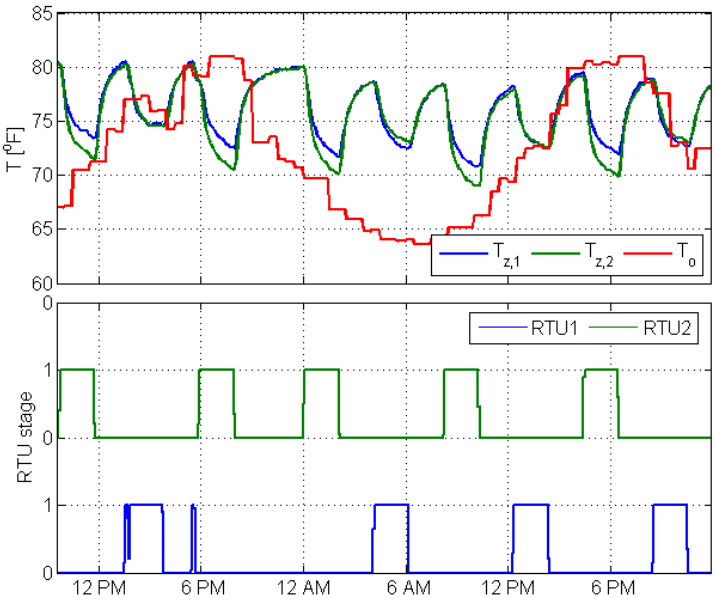


Figure 6: Responses of step tests for model comparisons and validation (Purdue Herrick Laboratory)

residual analysis does not answer the question.

A more convincing and straightforward model validation strategy is to find the "true system" and compare it to the two models. Therefore, in order to qualify the two models, we performed several step tests for the building and compared them to the step responses of the estimated models.

Note that it is impossible for a real building to reach an equilibrium point because uncontrollable inputs, e.g. weather and internal heat gains, change continuously and the building dynamics are typically slow because of high thermal mass of the building materials. Therefore, the experiment does not show the true step response in a strict sense. Nevertheless, we chose this approach because it at least provides some useful information of the building dynamics and there are no simple alternatives for model validation for a real building system.

To implement the step tests, all RTUs were forced to be off for about 2 hours in order to derive the state of the building dynamics at an "equilibrium" point. Then, we activated RTU1 for about 2 hours. After that, the RTU was turned off in order to allow the building dynamics return to the equilibrium point. The procedure was repeated for the other RTU. Fig. 6 shows the implementation and the responses of the zone air temperature. The outdoor air temperature during the experiment is also included in the figure.

For each period of the step tests, the temperature responses were subtracted from the equilibrium point, denoted as $\Delta T(k)$. The experimental step responses

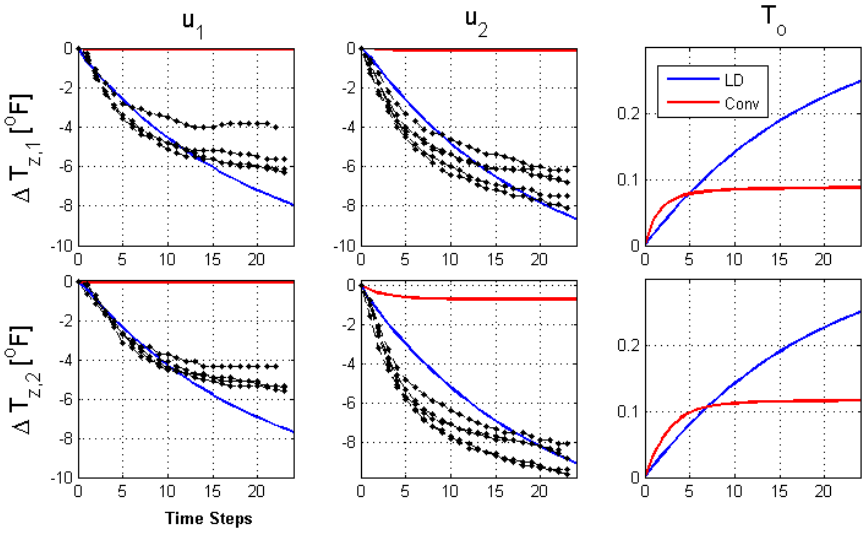


Figure 7: Model comparisons with experimental step test results (dotted lines), 1 time step = 5 min

of ΔT over the 2 hours are shown with dots in Fig. 7, where the 1 step in the x-axis represents 5 min. The experimental response do not match each other due to time varying outdoor air temperatures and unknown disturbances. Since the outdoor air temperature is not controllable, no step responses associated with the outdoor air temperature change are included in the figure. The unit step responses of the LD and Conv were calculated purely using estimated models and are shown with blue and red lines, respectively.

Note that the step response of Conv is far away from the experimental step test results. The Conv predicts that RTU1 has no cooling effect and RTU2 has much smaller cooling effect than the measurements. The inconsistency corresponds to (19) and (20) in the frequency domain. The distortion is intuitively obvious because the Conv has to explain input/output relationships without disturbances.

In contrast, the estimated model from the LD matches experimental results reasonably well, despite combinations of input-to-output responses to be explained. The inconsistencies could be from the narrow setpoint perturbation band, i.e. $\pm 2^o F$, or from the simple disturbance model structure. Nonetheless, the suggested approach alleviates influences of unmeasured disturbances on estimated models, and leads to significantly improved dynamic models for the MIMO building thermal system. The improvement from the Conv baseline is purely due to the additional disturbance model structure of the LD as shown in (25), because the same grey-box model structure, i.e. (16), and the same estimation data set, i.e. Fig. 5, were applied to both the LD and Conv.

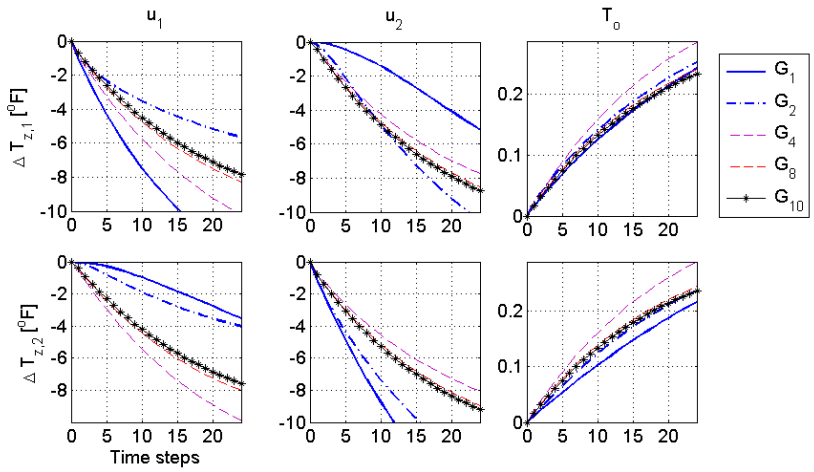


Figure 8: Behaviors of the LD w.r.t. sizes of estimation data sets, 1 time step = 5 min

More precisely, it is achieved by lowering the $\tilde{H}$-term in (11) and (12), which is the motivation of the LD approach, since the correlation terms in those equations are the same for both modeling methods. It should be noted that the reduced $\tilde{H}$ implies that the suggested output disturbance model structure (22) can efficiently capture complex building disturbances with a small number of parameters in this experiment.

Since T_o is not controllable, the step test approach is limited to validate estimated transfer functions associated with u_1 and u_2. However with the theoretical result in Section 2.4, i.e. an improvement of a model implies an improvement of the others, we conclude that the transfer function estimation for $G_{T,o}$ of the LD should be better than the Conv.

It is of interest to investigate the behavior of the LD with respect to a finite size of data set. Recall that the mathematical descriptions of (11) and (12) are valid for a large enough data set. For this purpose, the LD was applied to a sequence of subsets of the estimation data shown in Fig. 5: each subset data has a n-day data period from Aug/18 and n varies from 1 to 10 days. The resulting model corresponding to a n-day data set is denoted as G_n. Fig. 8 shows the behaviors of the estimated models as the data size increases. For a short training data period, e.g. less than 4 days, model performance is sensitive to the variation of data period. However, G_n starts to converge at every time step for every input/output pair after around one week showing that the LD is no longer sensitive to the data size.

4. CONCLUSIONS AND FUTURE WORK

This paper presents a unique mathematical framework in the building science field that provides qualitative guidelines for estimating the quality of building models that are trained using data obtained under closed-loop operation where there are significant unmeasured disturbances. This paper also presents a new identification approach based on the analytic result to extract an improved RC network building model from data under the presence of unknown heat gains. To reasonably characterize complex buildings disturbances, they are indirectly treated as output disturbances and are independently modeled from the semi-physical model. Using experiments, it was demonstrated that the method leads to a significantly improved model compared to that from a conventional grey-box simulation error minimization scheme. The results presented in this paper are originated from literature on system identification in other fields and are expected to be generally applicable to many different types of models for buildings.

5. APPENDIX

This section drives an expression for the asymptotic bias term for a general system shown in Fig. 1b.

Let $y(k) = \mathbf{G}_T(z)\mathbf{u}(k) + v(k)$ where $\mathbf{G}_T(z) = \begin{bmatrix} G_2(z) & G_1(z) \end{bmatrix}$ and $\mathbf{u}(k) = \begin{bmatrix} u_2(k) \\ u_1(k) \end{bmatrix}$.

The asymptotic objective function of a PEM for an MIMO system can be expressed as

$$\min_{\theta} trE(\epsilon(k)\epsilon^T(k)) \quad = \quad \min_{\theta} \frac{1}{2\pi} tr \int_0^{2\pi} \Phi_\epsilon(w)dw. \tag{27}$$

and

$$\Phi_\epsilon = H_\theta^{-1} \begin{bmatrix} \tilde{\mathbf{G}} & \tilde{H} \end{bmatrix} \begin{bmatrix} \Phi_{\mathbf{u}} & \Phi_{\mathbf{u}e} \\ \Phi_{e\mathbf{u}} & \Phi_e \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{G}}^\dagger \\ \tilde{H}^\dagger \end{bmatrix} H_\theta^{-\dagger} + \Phi_e. \tag{28}$$

where $\tilde{\mathbf{G}} = \mathbf{G}_T - \mathbf{G}_\theta$, $\tilde{H} = H_T - H_\theta$ and $\dagger$ represents conjugate transpose.

We want to decompose the expression of Φ_ϵ of (28) like (6).

Taking LDU decomposition on the power spectrum matrix in (28) leads to

$$\begin{bmatrix} \Phi_{\mathbf{u}} & \Phi_{\mathbf{u}e} \\ \Phi_{e\mathbf{u}} & \Phi_e \end{bmatrix} = \begin{bmatrix} I & \\ \Phi_{e\mathbf{u}}\Phi_{\mathbf{u}}^{-1} & I \end{bmatrix} \begin{bmatrix} \Phi_{\mathbf{u}} & \\ & \Phi_e - \Phi_{e\mathbf{u}}\Phi_{\mathbf{u}}^{-1}\Phi_{\mathbf{u}e} \end{bmatrix} \begin{bmatrix} I & \Phi_{\mathbf{u}}^{-1}\Phi_{\mathbf{u}e} \\ & I \end{bmatrix}. \tag{29}$$

Substituting (29) to (28) results in

$$\begin{aligned} \Phi_\epsilon - \Phi_e \quad &= \quad H_\theta^{-1}[(\tilde{\mathbf{G}} + \tilde{H}\Phi_{e\mathbf{u}}\Phi_{\mathbf{u}}^{-1})\Phi_{\mathbf{u}}(\tilde{\mathbf{G}}^\dagger + \Phi_{\mathbf{u}}^{-1}\Phi_{\mathbf{u}e}\tilde{H}^\dagger) \\ &+ \quad \tilde{H}(\Phi_e - \Phi_{e\mathbf{u}}\Phi_{\mathbf{u}}^{-1}\Phi_{\mathbf{u}e})\tilde{H}^\dagger]H_\theta^{-\dagger}. \end{aligned} \tag{30}$$

By introducing

$$\begin{aligned} B &:= \tilde{H}\Phi_{e\mathbf{u}}\Phi_{\mathbf{u}}^{-1} \\ \Phi'_e &:= \Phi_e - \Phi_{e\mathbf{u}}\Phi_{\mathbf{u}}^{-1}\Phi_{\mathbf{u}e}, \end{aligned} \tag{31}$$

(30) can be simply expressed as

$$\Phi_\epsilon - \Phi_e = H_\theta^{-1}(\underline{(\tilde{\mathbf{G}}+B)\Phi_{\mathbf{u}}(\tilde{\mathbf{G}}+B)^\dagger} + \tilde{H}\Phi'_e\tilde{H}^\dagger)H_\theta^{-\dagger}. \tag{32}$$

The equation (32) is an MIMO version of (6).

We want to split the expression (32) for bias terms corresponding to G_1 and G_2.
Since

$$\Phi_{\mathbf{u}} := \begin{bmatrix} \Phi_2 & \Phi_{21} \\ \Phi_{12} & \Phi_1 \end{bmatrix}, \tag{33}$$

B in (31) can be expressed as

$$B = \tilde{H}\begin{bmatrix} \Phi_{e2} & \Phi_{e1} \end{bmatrix}\begin{bmatrix} \Phi_2 & \Phi_{21} \\ \Phi_{12} & \Phi_1 \end{bmatrix}^{-1}. \tag{34}$$

Applying the matrix inversion lemma leads to

$$B = \tilde{H}\begin{bmatrix} \Phi_{e2} & \Phi_{e1} \end{bmatrix}\begin{bmatrix} \Phi_2^{-1} + \Phi_2^{-1}\Phi_{21}\Delta\Phi_{12}\Phi_2^{-1} & -\Phi_2^{-1}\Phi_{21}\Delta \\ -\Delta\Phi_{12}\Phi_2^{-1} & \Delta \end{bmatrix},$$

where $\Delta = (\Phi_1 - \Phi_{12}\Phi_2^{-1}\Phi_{21})^{-1}$.
After simple matrix manipulations, one can get

$$\begin{aligned} B &= \begin{bmatrix} B_2 & B_1 \end{bmatrix}, \text{ where} \\ B_2 &= \tilde{H}[\Phi_{e2}(\Phi_2^{-1} + \Phi_2^{-1}\Phi_{21}\Delta\Phi_{12}\Phi_2^{-1}) - \Phi_{e1}\Delta\Phi_{12}\Phi_2^{-1}] \\ B_1 &= \tilde{H}(\Phi_{e1} - \Phi_{e2}\Phi_2^{-1}\Phi_{21})\Delta. \end{aligned} \tag{35}$$

Therefore, the underlined term of (32) becomes

$$\begin{aligned} &(\tilde{\mathbf{G}}+B)\Phi_{\mathbf{u}}(\tilde{\mathbf{G}}+B)^\dagger \\ &= \begin{bmatrix} \tilde{G}_2 + B_2 & \tilde{G}_1 + B_1 \end{bmatrix}\begin{bmatrix} \Phi_2 & \Phi_{21} \\ \Phi_{12} & \Phi_1 \end{bmatrix}\begin{bmatrix} (\tilde{G}_2 + B_2)^\dagger \\ (\tilde{G}_1 + B_1)^\dagger \end{bmatrix}. \end{aligned} \tag{36}$$

One can arrive at the following expression after decomposing the power spectrum matrix in (36) as in (29).

$$\begin{aligned} &(\tilde{\mathbf{G}}+B)\Phi_{\mathbf{u}}(\tilde{\mathbf{G}}+B)^\dagger \\ &= \underbrace{(\tilde{G}_2 + B_2 + (\tilde{G}_1 + B_1)\Phi_{12}\Phi_2^{-1})}\Phi_2(\tilde{G}_2 + B_2 + (\tilde{G}_1 + B_1)\Phi_{12}\Phi_2^{-1})^\dagger \\ &+(\tilde{G}_1 + B_1)(\Phi_1 - \Phi_{12}\Phi_2^{-1}\Phi_{21})(\tilde{G}_1 + B_1)^\dagger. \end{aligned} \tag{37}$$

By substituting (35) into B_2, the underbraced term of (37) becomes

$$\begin{aligned}
&\tilde{G_2} + B_2 + (\tilde{G_1} + B_1)\Phi_{12}\Phi_2^{-1} \\
&= \tilde{G_2} + \tilde{H}[\Phi_{e2}(\Phi_2^{-1} + \Phi_2^{-1}\Phi_{21}\Delta\Phi_{12}\Phi_2^{-1}) - \Phi_{e1}\Delta\Phi_{12}\Phi_2^{-1}] \\
&+(\tilde{G_1} + \tilde{H}(-\Phi_{e2}\Phi_2^{-1}\Phi_{21}\Delta + \Phi_{e1}\Delta))\Phi_{12}\Phi_2^{-1} \\
&= \tilde{G_2} + \tilde{G_1}\Phi_{12}\Phi_2^{-1} + \tilde{H}\Phi_{e2}\Phi_2^{-1}.
\end{aligned} \tag{38}$$

Gathering (32),(35),(37) and (38) leads to the final asymptotic objective function of PEM for an MIMO system.

$$\min_{\theta} \frac{1}{2\pi} tr \int_0^{2\pi} \Phi_\epsilon(w)dw \tag{39}$$

where

$$\begin{aligned}
\Phi_\epsilon &= H_\theta^{-1}[(\tilde{G_2} + \tilde{G_1}\Phi_{12}\Phi_2^{-1} + \tilde{H}\Phi_{e2}\Phi_2^{-1})\Phi_2(\tilde{G_2} + \tilde{G_1}\Phi_{12}\Phi_2^{-1} + \tilde{H}\Phi_{e2}\Phi_2^{-1})^\dagger \\
&+ (\tilde{G_1} + B_1)(\Phi_1 - \Phi_{12}\Phi_2^{-1}\Phi_{21})(\tilde{G_1} + B_1)^\dagger \\
&+ \tilde{H}\Phi'_e\tilde{H}^\dagger]H_\theta^{-\dagger} + \Phi_e,
\end{aligned} \tag{40}$$

where

$$B_1 = \tilde{H}(\Phi_{e1} - \Phi_{e2}\Phi_2^{-1}\Phi_{21})(\Phi_1 - \Phi_{12}\Phi_2^{-1}\Phi_{21})^{-1}, \tag{41}$$

and $\Phi'_e = \Phi_e - \Phi_{e\mathbf{u}}\Phi_{\mathbf{u}}^{-1}\Phi_{\mathbf{u}e}$.

Consider a simple case when $u_1(k), u_2(k), y(k) \in \mathbb{R}$, then (39) and (40) becomes

$$\min_{\theta} E(\epsilon^2(k)) = \min_{\theta} \frac{1}{2\pi} \int_0^{2\pi} \Phi_\epsilon(w)dw \tag{42}$$

where

$$\begin{aligned}
\Phi_\epsilon &= |\tilde{G_2} + \tilde{G_1}\Phi_{12}\Phi_2^{-1} + \tilde{H}\Phi_{e2}\Phi_2^{-1}|^2 \frac{\Phi_2}{|H_\theta|^2} \\
&+ |\tilde{G_1} + B_1|^2 \frac{(\Phi_1 - \Phi_{12}\Phi_2^{-1}\Phi_{21})}{|H_\theta|^2} \\
&+ |\tilde{H}|^2 \frac{\Phi'_e}{|H_\theta|^2} + \Phi_e.
\end{aligned}$$

6. ACKNOWLEDGMENTS

This work was supported by the Department of Energy through the Consortium for Building Energy Innovation (CBEI) and by the Center for High Performance Buildings (CHPB) at Purdue University.

7. REFERENCES

[1] S. Prívara, J. Širokỳ, L. Ferkl, J. Cigler, Model predictive control of a building heating system: The first experience, Energy and Buildings 43 (2) (2011) 564–572, URL http://www.sciencedirect.com/science/article/pii/S0378778810003749.

[2] Y. Ma, F. Borrelli, B. Hencey, B. Coffey, S. Bengea, P. Haves, Model predictive control for the operation of building cooling systems, in: American Control Conference (ACC), 2010, IEEE, 5106–5111, URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5530468, 2010.

[3] J. Braun, Reducing energy costs and peak electrical demand through optimal control of building thermal storage, ASHRAE transactions 96 (2) (1990) 876–888.

[4] J. E. Braun, K. W. Montgomery, N. Chaturvedi, Evaluating the performance of building thermal mass control strategies, HVAC&R Research 7 (4) (2001) 403–428.

[5] O. Mejri, E. P. Del Barrio, N. Ghrab-Morcos, Energy performance assessment of occupied buildings using model identification techniques, Energy and Buildings 43 (2) (2011) 285–299.

[6] S. R. West, J. K. Ward, J. Wall, Trial results from a model predictive control and optimisation system for commercial building HVAC, Energy and Buildings 72 (2014) 271–279.

[7] W. J. Cole, K. M. Powell, E. T. Hale, T. F. Edgar, Reduced-order residential home modeling for model predictive control, Energy and Buildings 74 (2014) 69–77.

[8] H. Huang, L. Chen, E. Hu, A new model predictive control scheme for energy and cost savings in commercial buildings: An airport terminal building case study, Building and Environment 89 (2015) 203–216.

[9] D. Rogers, M. Foster, C. Bingham, A recursive modelling technique applied to the model predictive control of fluid filled heat emitters, Building and Environment 62 (2013) 33–44.

[10] D. Kolokotsa, A. Pouliezos, G. Stavrakakis, C. Lazos, Predictive control techniques for energy and indoor environmental quality management in buildings, Building and Environment 44 (9) (2009) 1850–1863, URL http://www.sciencedirect.com/science/article/pii/S0360132308002990.

[11] J. Bloem, System identification applied to building performance data, Office for Official Publications of the European Communities, 1994.

[12] L. Ferkl, J. Širokỳ, Ceiling radiant cooling: Comparison of ARMAX and subspace identification modelling methods, Building and Environment 45 (1) (2010) 205–212.

[13] P. R. Armstrong, S. B. Leeb, L. K. Norford, Control with building mass-Part I: Thermal response model, TRANSACTIONS-AMERICAN SOCIETY OF HEATING REFRIGERATING AND AIR CONDITIONING ENGINEERS 112 (1) (2006) 449.

[14] S. Prívara, J. Širokỳ, L. Ferkl, J. Cigler, Model predictive control of a building heating system: The first experience, Energy and Buildings 43 (2) (2011) 564–572.

[15] T. Chen, Real-time predictive supervisory operation of building thermal systems with thermal mass, Energy and Buildings 33 (2) (2001) 141–150, URL http://www.sciencedirect.com/science/article/pii/S0378778800000785.

[16] T. Zakula, P. Armstrong, L. Norford, Modeling environment for model predictive control of buildings, Energy and Buildings 85 (2014) 549–559.

[17] J. Širokỳ, F. Oldewurtel, J. Cigler, S. Prívara, Experimental analysis of model predictive control for an energy efficient building heating system, Applied Energy 88 (9) (2011) 3079–3087.

[18] B. Dong, K. P. Lam, C. Neuman, Integrated building control based on occupant behavior pattern detection and local weather forecasting, in: Twelfth International IBPSA Conference. Sydney: IBPSA Australia, 14–17, 2011.

[19] K. K. Andersen, H. Madsen, L. H. Hansen, Modelling the heat dynamics of a building using stochastic differential equations, Energy and Buildings 31 (1) (2000) 13–24.

[20] G. Reynders, J. Diriken, D. Saelens, Quality of grey-box models and identified parameters as function of the accuracy of input and observation signals, Energy and Buildings 82 (2014) 263–274.

[21] I. Naveros, C. Ghiaus, Order selection of thermal models by frequency analysis of measurements for building energy efficiency estimation, Applied Energy 139 (2015) 230–244.

[22] H. Harb, N. Boyanov, L. Hernandez, R. Streblow, D. Müller, Development and validation of grey-box models for forecasting the thermal response of occupied buildings, Energy and Buildings 117 (2016) 199–207.

[23] G. Fraisse, C. Viardot, O. Lafabrie, G. Achard, Development of a simplified and accurate building model based on electrical analogy, Energy and buildings 34 (10) (2002) 1017–1031.

[24] X. Chen, Q. Wang, J. Srebric, Model predictive control for indoor thermal comfort and energy optimization using occupant feedback, Energy and Buildings 102 (2015) 357–369.

[25] D. Rogers, M. Foster, C. Bingham, Experimental investigation of a Recursive Modelling MPC system for space heating within an occupied domestic dwelling, Building and Environment 72 (2014) 356–367.

[26] S. Prívara, J. Cigler, Z. Váňa, F. Oldewurtel, C. Sagerschnig, E. Žáčeková, Building modeling as a crucial part for building predictive control, Energy and Buildings 56 (2013) 8–22.

[27] P. Radecki, B. Hencey, Online building thermal parameter estimation via unscented kalman filtering, in: American Control Conference (ACC), 2012, IEEE, 3056–3062, URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6315699, 2012.

[28] Y. Heo, V. M. Zavala, Gaussian process modeling for measurement and verification of building energy savings, Energy and Buildings 53 (2012) 7–18.

[29] F. M. Gray, M. Schmidt, Thermal building modelling using Gaussian processes, Energy and Buildings 119 (2016) 119–128.

[30] P. Bacher, H. Madsen, Identifying suitable models for the heat dynamics of buildings, Energy and Buildings 43 (7) (2011) 1511–1522.

[31] M. J. Jimenez, H. Madsen, Models for describing the thermal characteristics of building components, Building and Environment 43 (2) (2008) 152–162.

[32] J. Kelso, Buildings energy data book, US Dept. of Energy .

[33] Z. Liao, A. L. Dexter, An inferential model-based predictive control scheme for optimizing the operation of boilers in building space-heating systems, Control Systems Technology, IEEE Transactions on 18 (5) (2010) 1092–1102.

[34] T. Söderström, P. Stoica, System identification, Prentice-Hall, Inc., 1988.

[35] L. Ljung (Ed.), System Identification (2Nd Ed.): Theory for the User, Prentice Hall PTR, Upper Saddle River, NJ, USA, ISBN 0-13-656695-2, 1999.

[36] I. Gustavsson, L. Ljung, T. Söderström, Identification of processes in closed loopidentifiability and accuracy aspects, Automatica 13 (1) (1977) 59–75.

[37] G. C. Goodwin, K. S. Sin, Adaptive filtering prediction and control, Courier Dover Publications, 2013.

[38] D. Shook, C. Mohtadi, S. Shah, Identification for long-range predictive control, in: IEE Proceedings D (Control Theory and Applications), vol. 138, IET, 75–84, URL http://digital-library.theiet.org/content/journals/10.1049/ip-d.1991.0010, 1991.

[39] D. S. Shook, C. Mohtadi, S. L. Shah, A control-relevant identification strategy for GPC, Automatic Control, IEEE Transactions on 37 (7) (1992) 975–980, URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=148352.

[40] D. Kim, J. Cai, K. B. Ariyur, J. E. Braun, System identification for building thermal systems under the presence of unmeasured disturbances in closed loop operation: Lumped disturbance modeling approach, Building and Environment 107 (2016) 169–180.