Simultaneous identification of building dynamic model and disturbance using sparsity-promoting optimization

Tingting Zeng, Jonathan Brooks, and Prabir Barooah
University of Florida
Gainesville, Florida USA

***Abstract*— We propose a method for identifying thermal building models for HVAC control in the presence of large, unmeasured disturbances. In addition, the method also identifies the effects of those unmeasured disturbances on the output. Our method uses ℓ_1-regularization to encourage the derivative of the identified scaled disturbance to be sparse, the motivation of which is physically meaningful. We test our method using training data from both open-loop and closed-loop simulations. Results show that the identified model can accurately identify the transfer functions from flow rate and supply-air temperature to room temperature in both cases, even in the presence of large, unmeasured disturbances, which makes it valuable for MPC applications.**

I. Introduction

Accurate thermal building models are useful for a variety of reasons, ranging from increasing energy efficiency to providing ancillary service to the power grid [1–4]. In addition to thermal building models, effects of unmeasured disturbances are also valuable, e.g., to predict future load from past data [5]. There has been much work regarding methods to obtain thermal building models [6–10]. However, it was shown in [11] that conventional methods can fail in the presence of large, unknown disturbances, and HVAC systems are affected by large, unknown disturbances, especially the cooling load induced by the occupants—both directly from body heat and indirectly from lights and other equipment they use. Therefore, methods to estimate thermal dynamics in the presence of large disturbances—and identifying the effects of those disturbances—are needed, which is the focus of this work.

Reference [11] sought to address this problem by identifying a lumped-disturbance (LD) model that is driven by white noise, where the input consists of outside temperature, solar irradiance, and HVAC power. Because the driving white noise, $\epsilon[k]$, is an unknown input (unmeasured disturbance), a prediction error method cannot be applied directly. Instead, the authors minimize $\sum \epsilon[k]^2$, noting that $\epsilon[k]$ is also an innovation sequence. The authors test the LD method using closed-loop training data from a building simulation, and it is shown that the accuracy of the LD method far exceeds that of a conventional prediction error method. The authors also employ the method in an actual building, and results appear consistent with those from simulation.

In [12], the authors propose a method to estimate model parameters for a commercial building and the corresponding unmeasured internal gain. This is achieved by separating the identification into two steps. First, it is assumed that the unmeasured internal gain is 0 during the weekends, and weekend data are used to obtain an estimate of the model parameters; i.e., weekend data provide a "clear" estimate of the model parameters, unperturbed by any unmeasured internal gain. Once the model parameters are obtained, using separate training data, the prediction error sequence is used to compute the internal gain (i.e., by inverting the dynamic equations to solve for the corresponding internal gain).

A recent paper [13] proposed a method to identify a model with cooling energy consumption as output and lighting/equipment schedule as an input; the equipment schedule is a way of providing a rough estimate of expected disturbances (e.g., higher consumption on weekdays). A Kalman filter is used to increase accuracy and provide on-line predictions of future energy consumption. The method is adaptive so that the model can change during "special operation situations," but this results in reduced prediction accuracy for longer prediction horizons. Additionally, the effects of the disturbance on the cooling load are not identified directly.

In the work presented here, we propose a method to identify building thermal dynamics from input/output data in the presence of large, unmeasured disturbances. The proposed method also produces an estimate of the effects of the unmeasured disturbance on the indoor air temperature. We solve a least-squares (LS) problem with an additional ℓ_1 penalty that encourages the derivative of the identified output disturbance to be sparse [14, 15]. The motivation for this is that large unmeasured disturbances, such as internal load due to occupants, are often piecewise-constant, which results in a sparse derivative; e.g., a large group of students enters a building for a class all at once and then leaves all at once when the class ends. By doing this, the motivation for our

This research is partially supported by NSF grants 1463316 and 1646229.

method is physically meaningful. We test our method via simulation, and results indicate that the method can accurately estimate both the thermal dynamics and the effects of unmeasured disturbances for open-loop data. With closed-loop data, the proposed method still identifies some of the transfer functions accurately but not as accurately as with open-loop data.

Neither the LD method of [11] nor the method proposed in [12] considers sparsity of the disturbance derivative. Instead, the identified disturbances in those works are effectively the signals needed to cover the gap between the predicted and measured outputs. Additionally, the method proposed in [12] requires training data when the unmeasured disturbance is small (such as a weekend) to first identify the thermal dynamics; in contrast, the method we proposed here does not require constraints on the unmeasured disturbance during the training period to identify accurate model parameters.

Our work is most closely related to [11]. Both our work and [11] are based on a standard resistor-capacitor (RC) thermal-network model, and the LD method in [11] directly identifies the thermal resistances and capacitances from the RC model, whereas our method identifies lumped combinations of those values. However, in the LD method, the assumed model structure combines flow rate and supply-air temperature into a single, lumped input—HVAC power. Conversely, we use flow rate and supply-air temperature as separate inputs for our identified model structure, which increases the difficulty of the identification because more variables must be identified. This also makes our method more readily applicable to many MPC formulations for smart buildings [3, 16, 17], since supply-air flow rate and supply-air temperatures can be individually commanded through variable-air-volume box dampers and reheat valves but HVAC power is a function of these two variables.

The rest of this paper is organized as follows. In Section II, we describe the model used for simulations to generate data. Section III formally describes the problem we need to solve and introduces the algorithms. We provide the evaluation results in Section IV. Finally, Section V concludes this work.

II. Building Model for Simulation

In order to evaluate the method, data need to be collected from a building model. Our starting point is a simple two-state RC-network nonlinear model [18], where the building is lumped into a single zone—we are after a lumped-parameter model with lumped disturbances acting on it. Figure 1 illustrates the lumped RC-network model.

The dynamics of the building's indoor temperature are affected by four *known* inputs: (1) the ambient

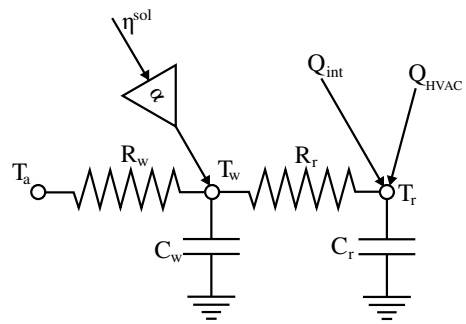

Fig. 1: A schematic of the nonlinear 2R2C model that motivates the LTI model structure used for simulation and identification. Q_{HVAC} represents the heat added or removed by the HVAC system.

(outside) temperature (T_a), (2,3) the mass flow rate and temperature of air supplied to the zone ($\dot{m}$,T_s), and (4) the solar irradiance (η^{sol}); and one unknown disturbance: the internal heat gain (Q_{int}) which comes from occupants, lights, and other equipment used by the occupants. The two states are the temperature of the room (T_r, the measurable output) and the temperature of the wall (T_w). The only measurable output is the indoor temperature. We then have

$$\begin{aligned} C_r\dot{T}_r &= \frac{1}{R_r}T_w - (\frac{1}{R_r} + \dot{m}C_{pa})T_r + \dot{m}C_{pa}T_s + Q_{int} \\ C_w\dot{T}_w &= \frac{1}{R_r}T_r - (\frac{1}{R_r} + \frac{1}{R_w})T_w + \frac{1}{R_w}T_a + \alpha\eta^{sol}, \end{aligned} \tag{1}$$

where α is solar absorbance ratio, $[C_r, C_w, R_r, R_w]^T$ are thermal capacitances and resistances of the room and wall, respectively.

Since we are aiming to identify model structure for each input separately and distinctly, we linearize (1).

A. Linearized RC Model

Define the states as $x_1 = T_r$ and $x_2 = T_w$ so that the state vector is $x = [T_r, T_w]^T$. Further, define $u = [\dot{m}, T_s, T_a, \eta^{sol}]^T$ and $w = Q_{int}$. Eq. (1) can now be expressed as:

$$\dot{x} = f(x, u, w)$$

Denote the equilibrium points by a star and the deviation states/inputs by a tilde so that

$$\dot{x}^* = f(x^*, u^*, w^*) = 0$$

$$\tilde{x} = x - x^*, \quad \tilde{u} = u - u^*, \quad \tilde{w} = w - w^*$$

The linearized, deviation model can be written as:

$$\begin{aligned} \dot{\tilde{x}} &= F\tilde{x} + G\tilde{u} + J\tilde{w}, \\ y &= H\tilde{x}, \end{aligned} \tag{2}$$

where $F = \frac{\partial f}{\partial x}|_{(x^*,u^*,w^*)}$, $G = \frac{\partial f}{\partial u}|_{(x^*,u^*,w^*)}$, $J = \frac{\partial f}{\partial w}|_{(x^*,u^*,w^*)}$, and $H = [1\ 0]$. This yields

$$F = \begin{bmatrix} -\frac{1}{C_r}(\frac{1}{R_r} + C_{pa}\dot{m}^*) & \frac{1}{C_r R_r} \\ \frac{1}{C_w R_r} & -\frac{1}{C_w}(\frac{1}{R_r} + \frac{1}{R_w}) \end{bmatrix},$$

$$G = \begin{bmatrix} \frac{C_{pa}}{C_r}(T_s^* - T_r^*) & \frac{\dot{m}^* C_{pa}}{C_r} & 0 & 0 \\ 0 & 0 & \frac{1}{R_w C_w} & \frac{\alpha}{C_w} \end{bmatrix},$$

$$J = \begin{bmatrix} \frac{1}{C_r} \\ 0 \end{bmatrix}.$$

III. Problem Formulation

A. Model structure for identification

By taking the Laplace transform of (2), we obtain

$$Y(s) = \frac{1}{\text{den}(s)} \left[s + \tfrac{1}{C_w}(\tfrac{1}{R_r} + \tfrac{1}{R_w}), \quad \tfrac{1}{C_r R_r} \right] \left(\begin{bmatrix} g_{11}\tilde{u}_1(s) + g_{12}\tilde{u}_2(s) \\ g_{23}\tilde{u}_3(s) + g_{24}\tilde{u}_4(s) \end{bmatrix} + \begin{bmatrix} \frac{1}{C_r}\tilde{w}(s) \\ 0 \end{bmatrix} \right), \tag{3}$$

where the g_{ij}'s are the entries of the matrix G and

$$\begin{aligned} \text{den}(s) &= \det(sI - F) \\ &= s^2 + As + B, \end{aligned} \tag{4}$$

with

$$A = \frac{1}{C_w}\Big(\frac{1}{R_r} + \frac{1}{R_w}\Big) + \frac{1}{C_r}\Big(\frac{1}{R_r} + \dot{m}^* C_{pa}\Big),$$

$$B = \frac{\dot{m}^* C_{pa}}{C_w C_r}\Big(\frac{1}{R_r} + \frac{1}{R_w}\Big) + \frac{1}{R_r R_w}.$$

Motivated by the continuous-time input-output model (3), we assume the following form of a discrete-time input-output model for identification, which is a discrete-time analogue of (3):

$$\tilde{T}_r(z) = \frac{1}{D(z)} K^T \begin{bmatrix} \tilde{\dot{m}}(z) \\ \tilde{T}_s(z) \\ \tilde{T}_a(z) \\ \tilde{Q}_s(z) \end{bmatrix} + \frac{\bar{w}(z)}{D(z)}, \tag{5}$$

where

$$K^T = \begin{bmatrix} \theta_3 z^{-2} + \theta_5 z^{-1} + \theta_7 \\ \theta_4 z^{-2} + \theta_6 z^{-1} + \theta_8 \\ \theta_9 z^{-2} + \theta_{10} z^{-1} + \theta_{11} \\ \theta_{12} z^{-2} + \theta_{13} z^{-1} + \theta_{14} \end{bmatrix}^T \tag{6}$$

and

$$D(z) = -\theta_2 z^{-2} - \theta_1 z^{-1} + 1. \tag{7}$$

The disturbance, $\bar{w}(z)$, represents a scaled transformation of the original input disturbance, $\tilde{w}$. The vector of parameters to be identified is denoted as

$$\theta := \left[\theta_1, \theta_2, \ldots \theta_{14}, \bar{\bar{\mathbf{w}}}\right] \in \mathbb{R}^{14+(N-2)},$$

where

$$\bar{\bar{\mathbf{w}}} := \left[\bar{w}[3], \bar{w}[4], \ldots, \bar{w}[N]\right]^T$$

is the scaled disturbance to be recovered and $\bar{w}[k]$ is the corresponding analogue of $\bar{w}(z)$ in discrete time.

B. Regressor form

Performing an inverse z-transformation on (5) and (6), yields a difference equation, from which we obtain the linear regression form:

$$y[k] = \phi[k]^T \theta, \quad k = 3, \ldots, N, \tag{8}$$

where the regressor is

$$\begin{aligned} \phi[k]^T := \Big[& y[k-1], y[k-2], u_1[k-2], \\ & u_2[k-2], \ldots, u_1[k], u_2[k], \\ & u_3[k-2], \ldots, u_4[k], e_{k-2}^T \Big] \in \mathbb{R}^{14+(N-2)}; \end{aligned}$$

e_k is the k-th canonical basis vector of $\mathbb{R}^{N-2}$:

$$e_k := [0, \ldots, 0, 1, 0, \ldots, 0]^T \in \mathbb{R}^{N-2},$$

where the 1 appears in the kth place. By using all the data from $k = 1$ to $k = N$, from (8) we obtain $Y = \Phi^T \theta$, where $Y := [y[3], \ldots, y[N]]^T \in \mathbb{R}^{N-2}$ and

$$\Phi := \begin{bmatrix} \phi[3]^T \\ \ldots \\ \phi[N]^T \end{bmatrix} \in \mathbb{R}^{N-2 \times N+12}$$

C. The sparsity assumption

The occupant-induced load w should not only be piecewise-constant, but it should change infrequently. That is, we want the vector $\Delta\bar{\mathbf{w}} := [\bar{w}[4] - \bar{w}[3], \ldots, \bar{w}[N] - \bar{w}[N-1]]^T$ to be sparse (i.e., mostly zeros). This can be encouraged by adding an ℓ_1 penalty, $\|D_1 \bar{\bar{\mathbf{w}}}\|_1$ [14, 15], to the prediction error $\|Y - \Phi\theta\|_2^2$, where $D_1 \in \mathbb{R}^{N-3 \times N-2}$ is the *first-difference matrix*:

$$D_1 := \begin{bmatrix} -1 & 1 & 0 & \ldots & \ldots \\ 0 & -1 & 1 & 0 & \ldots \\ & & \ddots & \ddots & \\ \ldots & \ldots & 0 & -1 & 1 \end{bmatrix}.$$

D. Optimization problem

The cost function for the optimization problem is

$$J(\theta) = \|Y - \Phi\theta\|_2^2 + \lambda \|\bar{D}_1 \theta\|_1 \tag{9}$$

where λ is a weighting factor and $\bar{D}_1 := [0 \mid D_1]$ so that $\bar{D}_1 \theta = D_1 \bar{\bar{\mathbf{w}}}$.

1) Stability constraints: It is reasonable that the plant should be stable, which means, for the discrete-time transfer function, the poles should lie inside the unit circle. That is, for (7), we have

$$\begin{aligned} -1 < \theta_1 < 1, \\ -2 < \theta_2 < 2. \end{aligned} \tag{10}$$

2) DC-gain constraints: Applying the Tustin approximation [19] to (4), it can be shown that the steady state of the denominator (7) is

$$\lim_{z\to 1} D(z) = 1 - \theta_1 - \theta_2 > 0. \tag{11}$$

The sign of the DC gain of the transfer function from each input to output is known; i.e., the DC gain of $\frac{K^T(z)}{D(z)}$ from $\dot{m}$ to T_r must be negative, and from T_s, T_a, η^{sol} to T_r must each be positive. Hence, based on (7), the steady state of each element of K in (6) should satisfy:

$$\begin{aligned} \theta_3 + \theta_5 + \theta_7 &< 0, \\ -\theta_4 - \theta_6 - \theta_8 &< 0, \\ -\theta_9 - \theta_{10} - \theta_{11} &< 0, \\ -\theta_{12} - \theta_{13} - \theta_{14} &< 0. \end{aligned} \tag{12}$$

The plant parameters and disturbances are then estimated by solving the optimization problem:

$$\hat{\theta} = \arg\min_{\theta} J(\theta) \tag{13}$$

subject to (10)-(12).

As in any ℓ_1-regularized LS problem, the choice of λ affects performance [14, 15], and some heuristic must be used to choose it.

E. Proposed algorithm

Input: N samples of input $\tilde{u}[k]$ and output $\tilde{y}[k]$.
Output: estimates of plant parameters $\theta_1, \theta_2, \ldots, \theta_{14}$ and scaled disturbances $\tilde{\bar{w}}[3], \ldots, \tilde{\bar{w}}[N]$.

Algorithm 1

Step 1: Choose operating points as $\bar{y} = \text{mean}(y)$ and $\bar{u} = \text{mean}(u)$, and compute deviation variables as $\tilde{y} = y - \bar{y}$, $\tilde{u} = u - \bar{u}$.

Step 2: Choose λ, and solve optimization problem (13) subject to (10)-(12).

Step 3: Continue to repeat step 2 until the smallest λ is found such that the estimated scaled disturbance is piecewise-constant.

Algorithm 1 performs well when the disturbance is truly piecewise-constant but not otherwise. Since disturbances are not likely to be truly piecewise-constant, even if they are likely to be approximately so, this motivates the development of a second algorithm to improve performance when the disturbance is not piecewise-constant. For ease of exposition, we define a matrix that will be used in the algorithm:

$$\bar{\Phi} = \Phi \begin{bmatrix} 0 \\ I \end{bmatrix} \in \mathbb{R}^{N-2\times N+10}, \tag{14}$$

where 0 and I are zero and identity matrices, respectively, of appropriate sizes (i.e., $\bar{\Phi}$ is Φ with the first two columns removed).

Algorithm 2

Step 1: Do step 1 from *Algorithm 1*.

Step 2: Do step 2 from *Algorithm 1*.

Step 3: Continue to repeat step 2 until the smallest λ is found such that the estimated scaled disturbance is nearly constant, and hold the corresponding $\hat{\theta}_1, \hat{\theta}_2$.

Step 4: Do *Algorithm 1* but replacing (13) with

$$\hat{\bar{\theta}} = \arg\min_{\bar{\theta}} \|Y - \bar{\Phi}\bar{\theta}\|_2^2 + \lambda \|\bar{\bar{D}}_1 \bar{\theta}\|_1$$

where $\bar{\theta} := \left[\theta_3, \theta_4, \ldots \theta_{14}, \bar{\bar{\mathbf{w}}}\right]$, $\bar{\bar{D}}_1 := [0 \mid D_1]$ so that $\bar{\bar{D}}_1 \bar{\theta} = D_1 \bar{\bar{\mathbf{w}}}$, and $\bar{\Phi}$ is defined in (14).

Step 5: Set $\hat{\theta} = \left[\hat{\theta}_1, \hat{\theta}_2, \hat{\bar{\theta}}\right]$.

IV. Evaluation of Proposed Algorithms

The linearized RC model (5), described in Section II, is used for simulations to generate training and validation data. Parameters are chosen as $(C_r, C_w) = [27.2, 20.4]$ (kWh/K), $[R_r, R_w] = [2.99, 1.04]$ (K/kW) and $\alpha = 0.5$, with $\dot{m}^* = 2.3$ (kg/s), $T_s^* = 21.7$ °C, and $T_r^* = 22.2$ °C.

Open-loop training data generated from this model are shown in Figures 2 and 3. The input data for $\tilde{\dot{m}}$ and $\tilde{T}_s$ in Figure 3 were collected from Pugh Hall on the University of Florida campus; the solar irradiance data are from SolRad-Net for the southeastern United States, and ambient temperature data are historical records from Gainesville, FL (both with 5-minute sampling periods).

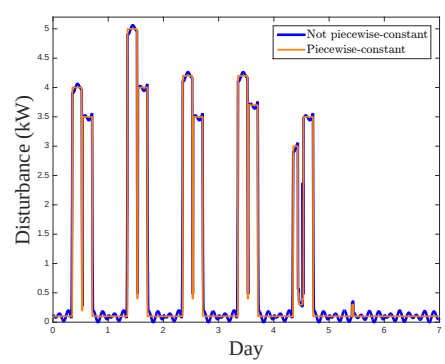

Fig. 2: Piecewise and not piecewise-constant disturbances for training and evaluation.

A. Case study 1: Piecewise-constant disturbance

For case study 1, we apply *Algorithm 1* on one day of training data with piecewise-constant disturbances (day 5 in Figures 2 and 3). Table I shows clearly that, with

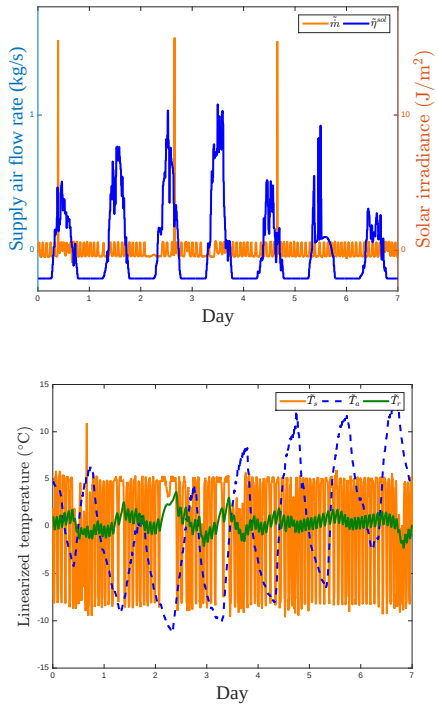

Fig. 3: Measurable inputs $[\dot{\tilde{m}}, \tilde{T}_s, \tilde{T}_a, \tilde{\eta}^{sol}]$ and output $\tilde{T}_r$ used for training.

Algorithm 1, the true system parameters can be identified accurately when the disturbance is piecewise-constant. Corresponding Bode plots are provided in Figure 4. The estimate of the scaled disturbance was also accurately identified, as shown in Figure 5.

TABLE I: Comparison of true system parameters and estimates for *Algorithm 1* (case study 1).

	Plant parameters θ	Estimated parameters $\hat{\theta}$
θ_1	1.95	1.95
θ_2	-9.52×10^{-1}	-9.52×10^{-1}
θ_3	6.29×10^{-3}	6.29×10^{-3}
θ_4	-1.26×10^{-2}	-1.26×10^{-2}
θ_5	-1.21×10^{-4}	-1.21×10^{-4}
θ_6	2.42×10^{-4}	2.42×10^{-4}
θ_7	-6.41×10^{-3}	-6.41×10^{-3}
θ_8	1.28×10^{-2}	1.28×10^{-2}
θ_9	1.27×10^{-5}	1.27×10^{-5}
θ_{10}	2.55×10^{-5}	2.55×10^{-5}
θ_{11}	1.27×10^{-5}	1.27×10^{-5}
θ_{12}	1.32×10^{-8}	1.33×10^{-8}
θ_{13}	2.65×10^{-8}	2.60×10^{-8}
θ_{14}	1.32×10^{-8}	1.35×10^{-8}

To further evaluate the estimate of θ, we use a validation data set to predict the room temperature. The validation data set again uses data from Pugh Hall for $\dot{\tilde{m}}$ and $\tilde{T}_s$ but from a different week; similarly, the solar irradiance and ambient air temperature were historical data from a different week than for the training data; however, the disturbance is the same (see Figure 2). We use the estimated $\hat{\theta}$ and the measured validation input data to predict the temperature in the validation data set, shown in Figure 6. We see that the model predicts the output extremely well.

B. Case study 2: Not piecewise-constant disturbance

Here, we apply *Algorithm 2*, with disturbance not piecewise-constant from Figure 2 and inputs from Figure 3 and with the corresponding output (datasets for a week). Again identification results are shown in Figures 4 and 7. Notice that for the inputs changing at a low frequency, i.e., T_a and η^{sol}, there exist minor errors because those data lack sufficient excitation at higher frequencies. However, estimates are still accurate at low frequencies, which is what will generally be encountered in practice. Again, we apply the estimated model to predict the room temperature from the validation data set; the predicted and actual temperatures are shown in Figure 6.

C. Case study 3: Closed-loop training data with piecewise-constant disturbance

While case studies 1 and 2 use open-loop training data, case study 3 uses closed-loop training data and piecewise-constant disturbance. A PI controller is implemented, where supply-air flow rate and supply-air temperature are adjusted to regulate the room temperature. The controller switches between two modes: heating (activated when $T_r < T_r^{ref}$ for longer than 5 minutes) and cooling (activated when $T_r > T_r^{ref}$ for longer than 5 minutes), where T_r^{ref} is the setpoint. For heating mode, $\tilde{T}_s \in [-0.05, 1.6]$ (° C) is determined by a PI controller with saturation, where proportional and integral gains are 0.03 and 0.001, respectively, and $\tilde{m} = -0.05$ (kg/s). Similarly, for cooling mode, $\tilde{m} \in [-10, 10]$ (kg/s) is determined by a PI controller with saturation, where proportional and integral gains are -0.03 and -0.001, respectively, and $\tilde{T}_s = -10$ (°C).

A closed-loop response was generated by adjusting the setpoint in order to get sufficiently exciting data. Disturbances are piecewise-constant, training data $(\tilde{T}_a, \tilde{\eta}^{sol})$ are the same as in case study 2. Training results for *Algorithm 2*, including Bode plots and scaled disturbance, are shown in Figures 9 and 10, respectively.

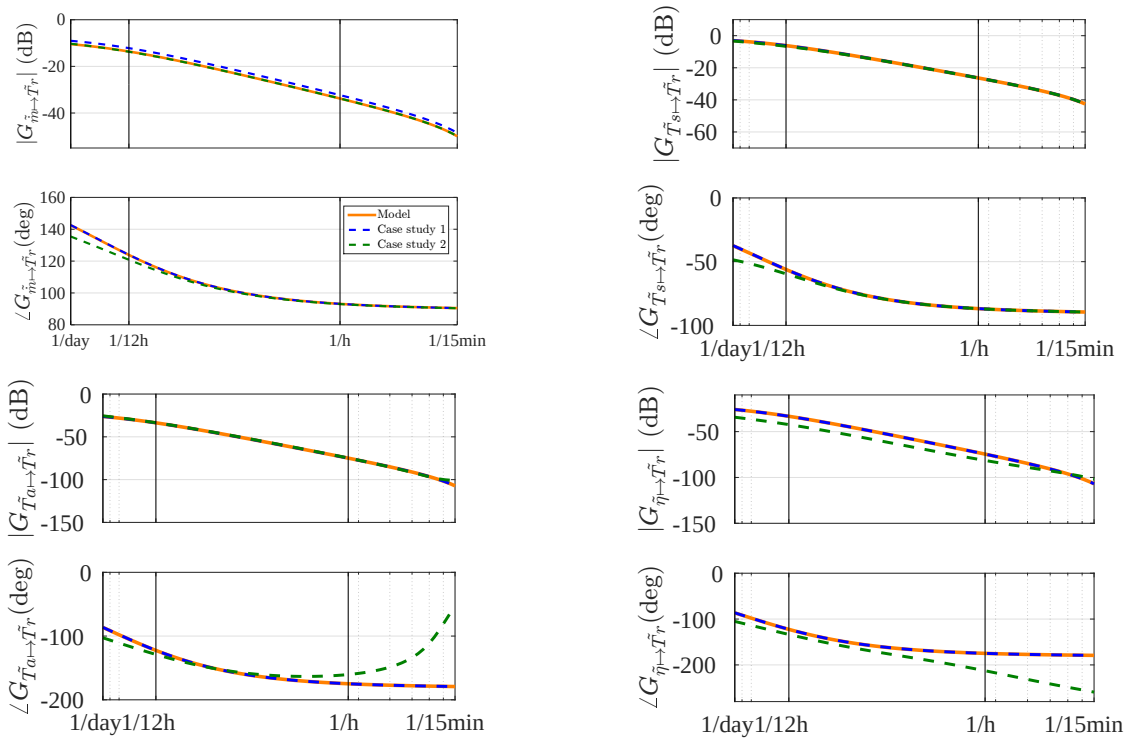

Fig. 4: Comparison of Bode plots of transfer function from each input to $\tilde{T}_r$ between true model and identified model for case studies 1 and 2.

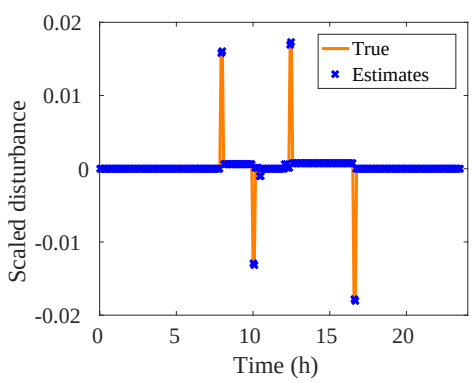

Fig. 5: Case study 1: Comparison of identified and actual scaled disturbance.

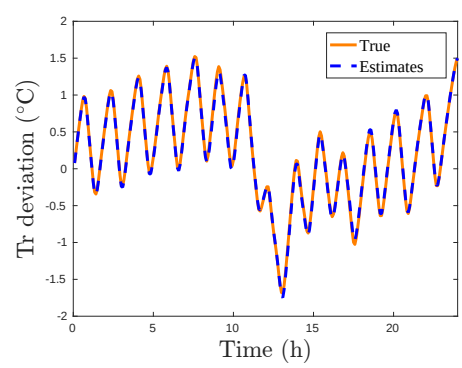

Fig. 6: Case study 1: Predicted and actual room temperature (validation datasets).

Again, errors occur in the transfer functions from the uncontrollable inputs to the output, which is inevitable because they are changing at low frequency without sufficient excitation. As a result, the scaled disturbance also shows inaccuracies.

For validation, we use the solar irradiance and ambient temperature from the validation data set; the temperature setpoint is changed arbitrarily every 30 minutes, and the disturbance is piecewise-constant (see Figure 2). Despite the inaccuracies in the estimated scaled disturbance, predictions of the room temperature, shown in Figure 11, are still accurate.

V. Conclusion

We presented a method to identify a thermal building model from input/output data in the presence of large, unmeasured disturbances, such as heat gain from occupants. The proposed method provides transfer functions from supply-air flow rate and supply-air temperature

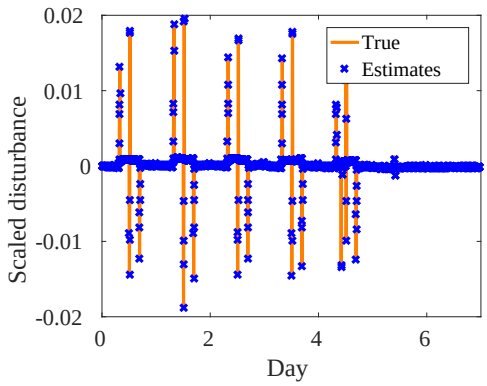

Fig. 7: Case study 2: Comparison of identified and actual scaled disturbance for a week.

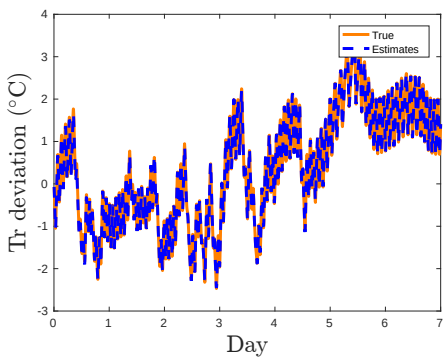

Fig. 8: Case study 2: Predicted and actual room temperature for a week (validation dataset).

to room temperature. The method also produces an estimate of the effects of the unmeasured disturbance. We encourage the estimate to be piecewise-constant through an ℓ_1 penalty on its derivative, which is a more physically meaningful constraint on the disturbance than in other works. For these reasons, the method is ideal for identifying models for the use of MPC.

We tested the proposed method using both open- and closed-loop training data from simulations. In the case of open-loop data, the proposed method accurately identified the model parameters and the effects of the disturbance. For closed-loop data, the method accurately identified the transfer functions from supply-air flow rate and supply-air temperature to room temperature, even in the presence of large, unmeasured disturbances. However, it struggled to accurately identify the effects of the unmeasured disturbance, due to lack of sufficient excitation of measured disturbances (solar irradiance and outside-air temperature) in the closed-loop training data. Even so, the transfer functions and scaled disturbances identified from closed-loop training data accurately predicted room temperature during validation.

Improving the proposed method's disturbance-identification for closed-loop training data is a subject of future work. Another avenue of future work is recovering individual resistance and capacitance values; currently, the method identifies parameters that are lumped combinations of these values. Finally, recovering the original input disturbance is also a subject of future work.

References

[1] F. Oldewurtel, A. Parisio, C. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and M. Morari, "Use of model predictive control and weather forecasts for energy efficient building climate control," *Energy and Buildings*, vol. 45, no. 2, pp. 15–27, February 2012.

[2] S. Prívara, J. Široký, L. Ferkl, and J. Cigler, "Model predictive control of a building heating system: The first experience," *Energy and Buildings*, vol. 43, no. 2-3, pp. 564–572, Feb. 2011. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0378778810003749

[3] Y. Ma, G. Anderson, and F. Borrelli, "A distributed predictive control approach to building temperature regulation," in *American Control Conference (ACC)*, July 2011.

[4] A. Afram and F. Janabi-Sharifi, "Theory and applications of HVAC control systems–a review of model predictive control (MPC)," *Building and Environment*, vol. 72, pp. 343–355, 2014.

[5] J. Page, D. Robinson, N. Morel, and J. Scartezzini, "A generalised stochastic model for the simulation of occupant presence," *Energy and Buildings*, vol. 40, no. 2, pp. 83–98, February 2008.

[6] J. E. Braun and N. Chaturvedia, "An inverse gray-box model for transient building load prediction," *HVAC&R Research*, vol. 8, pp. 73–99, 2002.

[7] S. Goyal, C. Liao, and P. Barooah, "Identification of multi-zone building thermal interaction model from data," in *50th IEEE Conference on Decision and Control and European Control Conference*, December 2011, pp. 181–186.

[8] 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.

[9] ——, "Building thermal model reduction using nonlinear constrained optimization," *Building and Environment*, vol. 37, no. 12, pp. 1255–1265, 2002.

[10] K.-h. Lee and J. E. Braun, "Development and application of an inverse building model for demand response in small commercial buildings," in *SimBuild 2004*. IBPSA-USA, 2004.

[11] D. Kim, J. Cai, K. B. Ariyur, and 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*, vol. 107, pp. 169 – 180, 2016.

[12] Q. Hu, F. Oldewurtel, M. Balandat, E. Vrettos, D. Zhou, and C. J. Tomlin, "Building model identification during regular operation-empirical results and challenges," in *American Control Conference (ACC), 2016*. IEEE, 2016, pp. 605–610.

[13] X. Li and J. Wen, "System identification and data fusion for on-line adaptive energy forecasting in virtual and real

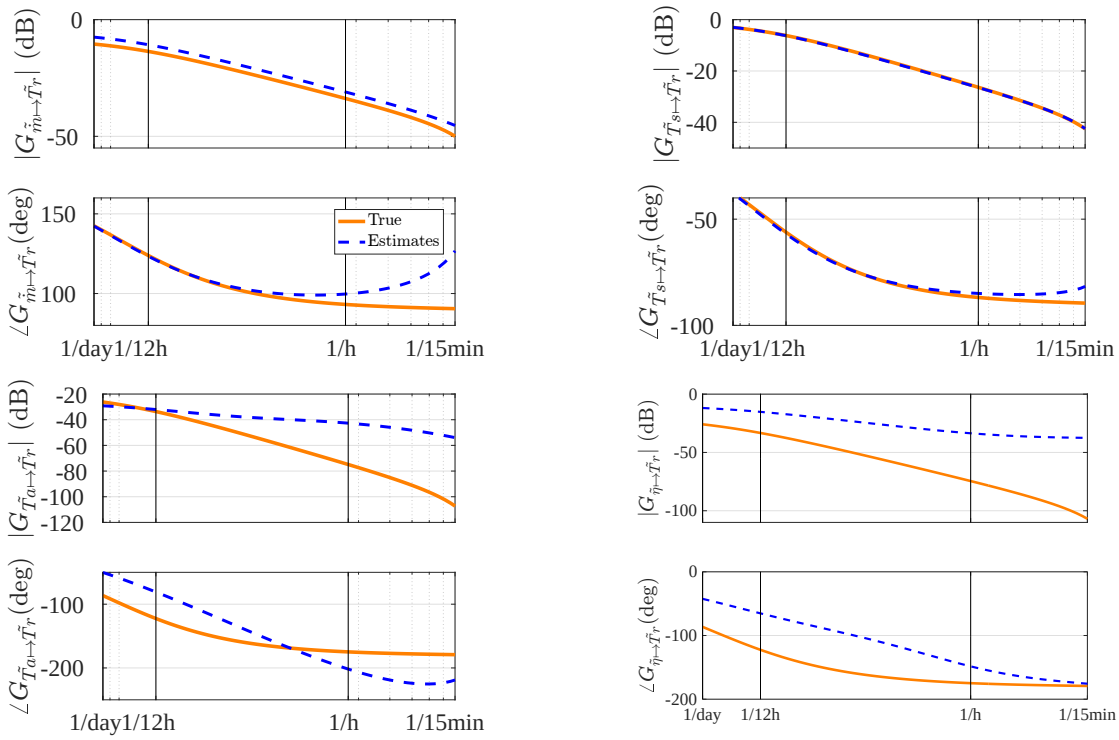

Fig. 9: Comparison of Bode plots between closed-loop plant and identified plant.

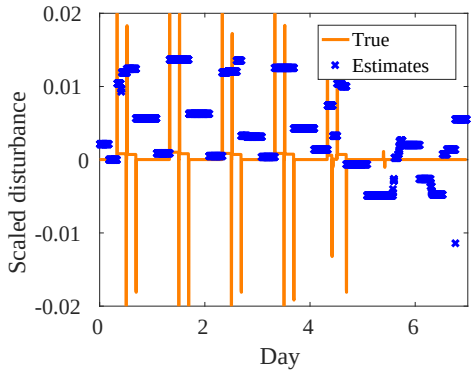

Fig. 10: Case study 3: Comparison of identified and actual scaled disturbance.

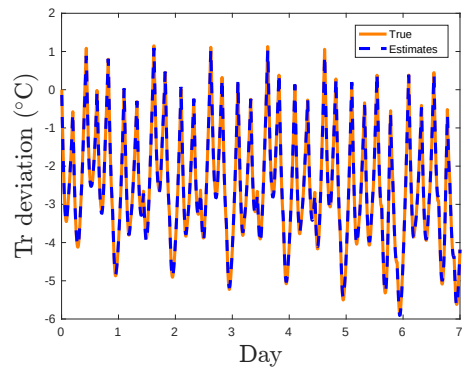

Fig. 11: Case study 3: Predicted and actual room temperature for a week (validation datasets).

commercial buildings," *Energy and Buildings*, vol. 129, pp. 227 – 237, 2016.

[14] D. M. Malioutov, M. Çetin, and A. S. Willsky, "Homotopy continuation for sparse signal representation," in *In Prceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2005, pp. 733–736.

[15] M. Ayazoglu, M. Sznaier, and N. Ozay, "Blind identification of sparse dynamic networks and applications," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*. IEEE, 2011, pp. 2944–2950.

[16] S. Goyal, P. Barooah, and T. Middelkoop, "Experimental study of occupancy-based control of HVAC zones," *Applied Energy*, vol. 140, p. 7584, February 2015.

[17] M. Maasoumy, M. Razmara, M. Shahbakhti, and A. S. Vincentelli, "Selecting building predictive control based on model uncertainty," in *American Control Conference (ACC), 2014*. IEEE, 2014, pp. 404–411.

[18] S. Goyal and P. Barooah, "A method for model-reduction of nonlinear building thermal dynamics," in *In proceedings of the 2011 American Control Conference*, July 2011, pp. 2077–2082.

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