Simultaneous identification of dynamic model and occupant-induced disturbance for commercial buildings[☆]

Austin R. Coffman*, Prabir Barooah

Mech. and Aero. Engg., Univ. of Florida, Gainesville, FL 32611, USA

Abstract

A model of a building's thermal dynamics is needed for prediction-based control. The task of identifying a thermal dynamic model is made challenging by the presence of large unmeasured disturbances, especially the heat gain due to the occupants. In fact, identification of this "occupant-induced load" is also valuable for predictive control—especially in commercial buildings. We propose a method to identify both a model (of resistance-capacitance network type) and the unmeasured disturbances from measured input-output data. The method is based on the insight that the main contributor to the unmeasured disturbance, the occupant-induced load, is piecewise constant, especially in commercial buildings. This can be used to construct an augmented dynamic model so that disturbance estimation is converted to a state estimation problem. An outer-loop optimization identifies the best-fit parameter values. The effectiveness of the method is evaluated using data from a simulation model (under both open and closed-loop operations) and a real building.

Keywords: system identification, HVAC control, disturbance modeling, building thermal dynamics, data driven modeling

1. Introduction

A dynamic model of a building's zone temperature is useful in several applications, particularly model-based control of the heating, ventilation, and air conditioning (HVAC) system for improving indoor air quality and reducing energy use [1], limiting peak demand [2], or providing ancillary services to the power grid [3]. To be useful in online control computations, the model needs to be of low order. To model a building with a small number of parameters, the model needs to be identified from data. That is, the parameters need to be estimated from measurements of inputs (such as ambient temperatures and control commands) and outputs (indoor temperature).

Among many different classes of models, a commonly used subclass is the resistance-capacitance (RC) networks, in which an electrical analogy is used to model heat flow. Many methods for estimating parameters of an RC network model from data have been proposed in the literature; see [4, 5, 6, 7] for early work, and [8, 9, 10, 11, 12, 13] for more recent work. In such a model, the output is the building's indoor temperature and the inputs consist of signals such as rate of cooling/heating provided by the HVAC system and dry bulb temperature of the air outside the building. Another input that affects the temperature is the "occupant-induced" heat-gain, which consists of the heat released by the occupants' bodies and the heat gain due to appliances used by the occupants. This input is difficult to measure and is therefore best modeled as an unmeasured disturbance.

With the notable exception of [10, 11, 13], other papers cited above either neglect the effect of unmeasured disturbance altogether [8, 9, 12], use an unoccupied test building [5, 6], or utilize previous knowledge of the occupancy schedule or specific instrumentation to make knowledge of the occupant-induced load possible [4, 7]. The last approach is limited to specially instrumented buildings. Neglecting disturbance altogether is a poor choice: [10] showed that doing so results in poor model identification. This is not surprising since the occupant-induced heat-gain, the main component of the unmeasured disturbance, is not small. Sometimes it can be comparable to the cooling provided by the HVAC system.

In this paper, we propose a method to simultaneously estimate (i) the parameters of an RC network model of a building zone and (ii) the unmeasured disturbance that acts as an exogenous input, from measurements of input-output data. We call it the Simultaneous Plant and Disturbance Identification (SPDI) algorithm; it is based on the assumption that the unmeasured disturbance is dominated by occupant-induced load, and the latter is piecewise constant so that its time-derivative is mostly zero.

[☆]The research reported here has been partially supported by NSF grants 1646229 (CPS-ECCS) and 1463316 (CMMI), and a DOE grant titled "Virtual Batteries" from the BTO's GMLC program.

*Corresponding author

Email addresses: bubbaroney@ufl.edu (Austin R. Coffman), pbarooah@ufl.edu (Prabir Barooah)

This enables us to model the unmeasured disturbance as a state of an augmented dynamic model, eliminating the unmeasured disturbance from the model. The insight that occupant-induced load is approximately piecewise constant is based on how commercial buildings are used. People come into a building in the morning so that the building goes from mostly empty to mostly full in a short period of time. The opposite happens in the evening, with some small variations during the working hours. For a given set of model parameters, the disturbance can be now estimated using a state estimator. An optimizer identifies the best-fit parameters. Apart from the R,C parameter values, the proposed method also identifies the effective area of the building that determines the total solar heat gain.

Identifying the unknown disturbance is useful in its own right for predictive control of HVAC systems. Once it is identified from past data, a model can be trained on that estimate which can then be used to predict the disturbance in the future. These predictions can be used by a control algorithm to make decisions for the HVAC system.

Application of the method to simulation-generated data, where the ground truth is known, shows that the method works well. It identifies the model and the disturbance with high accuracy even when the disturbance is not piecewise constant. In that case, the method is able to identify a smoothed version of the time-varying disturbance. The method's performance is insensitive to the use of data collected under normal (i.e., closed-loop) operation, which is usually a concern [8, 10, 11].

A challenge for the evaluation of building model identification methods is the lack of ground truth. In [10], evaluation of the model identified from real building data was performed by comparison of the model's step response to an experimentally determined step response of the building. This method is not applicable to evaluating the identified disturbance.

In this paper we evaluate the accuracy of the identified occupant-induced load by qualitatively comparing it with measured CO_2 concentration. In the building whose data the method was tested on, the control system increases outdoor airflow in response to an increase in measured CO_2 which occurs due to an increase in occupancy. If the disturbance q_{dist} is mostly induced by occupants, then it should be strongly correlated with CO_2 concentration. A high correlation between the estimated disturbance and the measured CO_2 concentration is an indication of the accuracy of the estimate.

The rest of the paper is organized as follows. Section 1.1 describes the contribution of this work with respect to prior art. Section 2 states the problem addressed in this paper precisely, and Section 3 describes the proposed algorithm. Sections 4 and 5 describe evaluation of the method when applied to simulation generated data and data collected from a real building, respectively. The paper ends with a conclusion in Section 6.

1.1. Contribution over prior art

To the best of our knowledge, the only references that deal with the problem of identifying both a RC network-type model and unmeasured disturbance from data are [10, 11, 13].

The method proposed by Kim *et al.* [10] does not identify the unmeasured input disturbance. Rather, an output disturbance (a disturbance that is added to the plant output) is identified that encapsulates the effect of an unknown input disturbance. The identified output disturbance is not a physically meaningful signal, though it could be potentially useful for predictive control. In contrast, the input disturbance we identify is a physically meaningful quantity as long as it is dominated by occupant-induced load. As a result, the accuracy of the identified disturbance by the proposed method can be qualitatively assessed by comparing with occupant schedules or measurable surrogates for the occupant-induced load such as CO_2 concentration [14] and WiFi client count [15].

Both [11] and [13] take a similar approach: the model is estimated by using data from unoccupied periods (weekends in [11]) and assuming that the disturbance is zero during those periods. Once the model is identified this way, the disturbance is identified using data from occupied periods. This approach has two weaknesses. First, it requires availability of sufficient training data collected when the building is unoccupied. Such periods may be short compared to the length of the data needed for training. Sometimes one may erroneously believe the building to be unoccupied. Second, by forcing the disturbance to be zero during unoccupied periods, the disturbance term is prevented from absorbing unmodeled dynamics—such as ground coupling and effect of latent heat exchange—that may be present even in unoccupied times. In contrast, the method proposed here does not require data from unoccupied periods. Data from any period can be used. In fact, the proposed method is used on data collected during regular operation of a building (see Section 5), while the data used in [11] were collected during an excitation test in which the air flow rates were temporarily set at the maximum.

The solar heat gain is assumed to be a known input in [13]. In practice, while solar irradiance can be easily measured with a sensor, there is significant uncertainty in the relationship between solar irradiance and solar heat gain. In the proposed method, we only use solar irradiance as a known input. The effective area, which determines the solar heat gain together with solar irradiance, is an unknown parameter to be estimated.

There are some similarities between our work and that of [13]. The assumption that internal loads are mostly constant, and therefore can be modeled as a process with a zero derivative, is used in both.

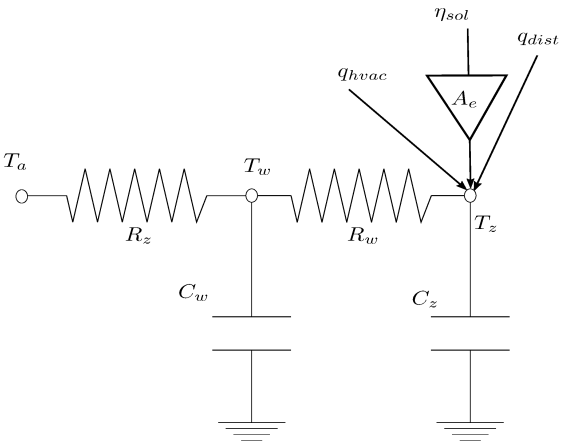

Figure 1: RC-network model structure for building thermal dynamics.

2. Problem statement and insight into proposed method

Figure 1 shows a schematic representation of the RC network model used in this paper, which has two capacitances C_z, C_w and two resistors R_z, R_w. The capacitance C_w (for "wall") is associated with the large thermal capacitance of the building's external structure while C_z (for "zone") is associated with the smaller thermal capacitance of the remaining objects in the building. Together they are likely to capture both the slow and fast time scales of temperature variation. The resistance R_w and R_z capture the resistance to heat flow between the ambient air and building thermal mass, and between the building thermal mass and zone air mass, respectively.

For large buildings with many zones that exhibit several distinct timescales of temperature variation, more complex models with more states and parameters may be needed. In this paper we focus on buildings for which two time scales are adequate.

The model itself is a coupled ordinary differential equation with two states, T_z (zone, or indoor, temperature) and T_w (wall temperature):

$$C_z\dot{T}_z = \frac{(T_w(t) - T_z(t))}{R_w} + q_{\text{hvac}}(t) + q_{\text{dist}}(t) + A_e\eta_{\text{sol}}(t) \quad (1)$$

$$C_w\dot{T}_w = \frac{(T_a(t) - T_w(t))}{R_z} + \frac{(T_z(t) - T_w(t))}{R_w} \quad (2)$$

where (i) heat gain due to the HVAC system q_{hvac} (kW), (ii) ambient (i.e., outdoor) temperature T_a (°C), and (iii) solar irradiance η_{sol} (kW/unit area), are measurable inputs, the indoor temperature, T_z, is a measurable output, and q_{dist} is the unmeasured disturbance that captures the occupant-induced internal heat load and any unmodeled internal dynamics. The term $A_e\eta_{\text{sol}}(t)$ is the heat gain due to solar radiation, where A_e is the effective area of the building. In general, all five parameters R_w, R_z, C_w, C_z and A_e are unknown for a given building.

The identification problem we address in this paper is the following: *given time traces of the measured inputs* $(q_{\text{hvac}}(t), T_a(t), \eta_{\text{sol}}(t))$ *and the output* $(T_z(t))$*, identify the unknown parameters* C_z, C_w, R_w, R_z, A_e *as well as the disturbance* $q_{\text{dist}}(t)$. Only sensible loads are considered in this work, the effects of humidity and ground coupling are ignored. The disturbance q_{dist} plays the role of collecting all the unmodeled dynamics.

It should be clear that without further information or constraints, this problem is inherently ill-posed: given the knowledge of input u and output y, it is not possible to determine the unknowns a and w uniquely from the relation $y = au + w$. To make the identification problem well-posed, we introduce the following additional constraint on the disturbance, namely, that it is piecewise constant. In other words, for almost all time t,

$$\frac{d}{dt}q_{\text{dist}}(t) = 0. \quad (3)$$

The motivation for this assumption comes from the hypothesis that the disturbance is mostly occupant induced load (equivalently, the contribution from unmodeled dynamics is small), and the fact that occupant-induced load varies in an approximately piecewise constant fashion in commercial buildings, in lockstep with how occupancy count changes over the day.

With the constraint (3), the signal q_{dist} becomes a state rather than an exogenous input. The augmented model is:

$$\begin{bmatrix}\dot{T}_z\\ \dot{T}_w\\ \dot{q}_{\text{dist}}\end{bmatrix} = \begin{bmatrix}\frac{-1}{R_wC_z} & \frac{1}{R_wC_z} & \frac{1}{C_z}\\ \frac{1}{R_wC_w} & \frac{-1}{C_w}(\frac{1}{R_w}+\frac{1}{R_z}) & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}T_z\\ T_w\\ q_{\text{dist}}\end{bmatrix} + \begin{bmatrix}0 & \frac{A_e}{C_z} & \frac{1}{C_z}\\ \frac{1}{R_zC_w} & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}T_a\\ \eta_{\text{sol}}\\ q_{\text{hvac}}\end{bmatrix} \quad (4)$$

$$T_z = \begin{bmatrix}1 & 0 & 0\end{bmatrix}\begin{bmatrix}T_z\\ T_w\\ q_{\text{dist}}\end{bmatrix}$$

In the sequel, the state matrix, the input matrix, and the output matrix in (4) will be referred to as A_c, B_c, C_c, respectively.

3. Proposed method: SPDI algorithm

If the parameters of the model were known, the problem of identifying the signal q_{dist} from input-output data becomes a state estimation problem. Our approach therefore relies on using a state estimator, in particular, a Kalman filter, for a given set of parameter values, and then searching for an optimal set of parameter values to minimize a cost function.

To enable use of sampled data, the continuous time dynamics (4) are discretized with a sampling period of

t_s to obtain a discrete-time linear time invariant (LTI) system [16]. Additionally, process and measurement noise are introduced to capture modeling error and sensor noise. The discrete-time model is therefore of the form

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + \xi_k \\ z_k &= Cx_k + n_k \end{aligned} \tag{5}$$

where ξ_k and n_k are white noise sequences that capture the modeling error and sensor noise, respectively, $A = e^{A_c t_s}$, $B = (\int_0^{t_s} e^{A_c \eta} d\eta) B_c$, and $C = C_c$. In the sequel, the process and measurement noise covariance sequences will be denoted by Q_k and R_k, i.e., $Q_k := Cov[\xi_k, \xi_k]$, $R_k := Cov[n_k, n_k]$, where $Cov(\cdot, \cdot)$ denotes covariance.

The Kalman filter uses the discrete-time system matrices (A, B, C), the measurement sequence $z_k := T_z[k] \in \mathbb{R}$, and the input sequence $u_k := [T_a[k], \eta_{\text{sol}}[k], q_{\text{hvac}}[k]]^T \in \mathbb{R}^3$, along with the covariances Q_k, R_k to produce the state estimate sequence $\hat{x}_{k|k} := E[x_k | z_0, \ldots, z_k]$, where $E[\cdot]$ denotes expectation. The residual sequence of the filter is defined as

$$r_k := z_k - \hat{z}_{k|k} = z_k - (C\hat{x}_{k|k}) \tag{6}$$

In order to use a Kalman filter to obtain the state estimates from input and output data, the system model, i.e., the matrices (A, B, C) must be provided. This is equivalent to providing the parameter vector $p = [C_z, C_w, R_w, R_z, A_e]^T \in \mathbb{R}^5$. A measure how good a given choice of p is

$$f(p) = \sum_{k=1}^{k_{max}} |r_k|, \tag{7}$$

where r_k is the residual sequence defined in (6) and k_{max} is the number of samples in the dataset. While the parameters do not explicitly show in the definition (6), they affect the residuals. Hence the objective function f is a function of p. The best-fit parameter values p^*, for a given data set $\{u_k, z_k\}_{k=1}^{k_{max}}$, is the one that minimizes this cost. That is,

$$p^* = \arg\min_{p \in \Omega} f(p), \tag{8}$$

where Ω is a subset of $\mathbb{R}^5$ where the parameter vector p is constrained to lie. The optimal p^* is the estimated parameters for a given initial parameter vector, and the corresponding state estimates from the Kalman filter also provides the disturbance estimate $\hat{q}_{dist}$.

The objective $f(p)$ is a non-convex function of the parameter vector p. Therefore, there may not be a unique minimum p^*, and the minimum that any numerical search algorithm converges to (when it does converge) may depend on the initial guess. The optimization problem (8) is solved multiple times for distinct initial guesses. The mode of these local optimal points is taken as the estimate.

The proposed method, which we call the Simultaneous Plant and Disturbance Identification (SPDI) algorithm, is summarized below:

SPDI Algorithm

Inputs: $\mathbf{u} = \{u_k\}_{k=1}^{k_{max}}$, $\mathbf{z} = \{z_k\}_{k=1}^{k_{max}}$, process and measurement noise covariance sequence $\{Q_k, R_k\}_{k=1}^{k_{max}}$, search algorithm and stopping criterion, method for generating an initials guess, and the number of initial guesses to try, n_{ig}.

IG Generate initial guess, $p_0^{(\ell)}$, iteration ℓ.

Begin Optimization solver loop, iteration j

step 1 Based on parameter estimate $p_j^{(\ell)}$, compute the discrete-time matrices (A, B, C) in (5).

step 2 Perform state estimation using the Kalman filter to obtain $\{\hat{x}_{k|k}\}_{k=1}^{k_{max}}$, and compute the objective function value (7).

step 3 Update the next parameter vector to try $p_{j+1}^{(\ell)}$, using a search technique.

step 4 Repeat steps 1-3 until stopping criterion is satisfied.

End Optimization solver loop

Store results: $p^{*(\ell)} = [C_z, C_w, R_w, R_z, A_e]^T$ and $q_{\text{dist}}^{*(\ell)}$.

Compute estimate Compute the mode of the optimal solutions $\{p^{*(\ell)}\}_{\ell=1}^{n_{ig}}$, and return it as the estimate $\hat{p}$. Let i be an index that achieves the mode. Return the corresponding load as the estimated load: $\hat{q}_{\text{dist}} := q_{\text{dist}}^{*(i)}$.

Any optimization algorithm for conducting the search in step 3 and any method of generating initial guesses can be used. In this paper we use the MATLAB© `fmincon` function with default options to search for the minimum, which uses an interior point method [17], and Latin hypercube sampling to generate an initial guess [18]. All numerical results presented later in this paper were obtained with $n_{ig} = 50$, which was picked somewhat arbitrarily.

3.1. Choosing design-variables

A number of choices have to be made by the user in implementing the algorithm: (i) the process and measurement error covariance sequences $\{Q_k, R_k\}$, (ii) the region Ω, and (iii) the length of the data sequence, k_{max}.

In the results reported in the rest of the paper, the process and noise covariances are chosen to be time invariant. In particular, $Q_k = Q$ and $R_k = R$ for all k, where

$$Q := diag(\alpha_1, \alpha_1, \alpha_2) \tag{9}$$

Where both α_2 and α_1 are small positive numbers with $\alpha_2 \gg \alpha_1$. This is equivalent to saying that there is very little modeling error in the first two states, but the dynamics of the third state has large uncertainty. The rationale for this choice is that even in commercial buildings the piecewise constant model for the disturbance state is not always accurate, especially during periods when people come in

and leave. So the third state has large uncertainty. In this paper, we choose $\alpha_2 \approx 10^{-3}$ and $\alpha_1 = 10^{-7}$.

The value of R is chosen to be on the same order of magnitude as α_2 (10^{-3}) but the ratio, $\frac{\alpha_2}{R}$, is elected to be greater than one. This is because it is desirable to ascribe the difference between measurements and predictions to the state associated with α_2 (i.e. the disturbance, q_{dist}). If the ratio is larger, the Kalman filter ascribes more of the prediction error r_k to the third state, the q_{dist}, while if the ratio is smaller, it ascribes the prediction error to measurement noise. In the simulation study presented in Section 4, the ratio is elected to be $\frac{\alpha_2}{R} = 3$ for both closed-loop and open-loop data. When the building is under closed-loop control, the HVAC system actively tries to maintain the zone temperature T_z by varying q_{hvac} in response to the error between T_z and its desired value, T_{ref} (reference signal). During open-loop operation no such feedback is used. If the method were to be applied on a different building, with a potentially different RC network model, the values α_2 and R may have to be re-calibrated.

The choice of the region Ω is somewhat arbitrary, and requires some a-priori knowledge. A rough estimate of the range of A_e can be obtained from knowledge of building geometry, which can then be used to set bounds. In this paper, we apply our method to data from a section of a large building, which is described in Section 5. The upper bound on the A_e is computed from the roof area of that zone. The values of resistances and capacitances that were identified for a 100 m^2 (floor area) $\times$ 4.5m (height) zone in a large commercial building in [10], expressed in K/kWh and kW/K, respectively, were of $O(1)$. Our simulation data is generated from a model calibrated to a 540 m^2 (floor area) $\times$ 6.7 m (height) commercial building. Thus, the following bounds on the parameters are used in this paper (units are shown in parenthesis):

$$0.01 \leq C_w, C_z \ (\frac{\text{K}}{\text{kWh}}), R_w, R_z \ (\frac{\text{kW}}{\text{K}}), A_e \ (\text{m}^2) \leq 50 \quad (10)$$

The length of data sequence should be enough so that (i) the transients due to initial conditions die out and do not effect the results significantly [4], and (ii) the data captures the difference between the weekday and weekend loads. We have found through trial and error that data for two weeks is adequate for this purpose.

3.2. Kalman filter performance

The question of whether the Kalman filter, for the true model, produces accurate state estimates or not, is of paramount importance for the effectiveness of the method. This question reduces to the question of stabilizability and detectability, or their less general counterparts, controllability and observability [19]. If $(A, Q^{1/2})$ is stabilizable and (A, C) is detectable, where Q is the covariance of the process noise ξ_k, then the covariance of the Kalman filter's state estimates converges to a positive definite constant matrix [20]. That provides confidence on the accuracy of the state estimates.

Since $C = [1,\ 0,\ 0]$, the observability matrix is $\mathcal{O} = [C^T,\ (CA)^T,\ (CA^2)^T]^T$. The determinant turns out to be

$$det(\mathcal{O}) = \frac{t_s^2(R_w t_s + R_z t_s)}{C_w C_z R_w^2 R_z}, \quad (11)$$

which is non-zero for physically meaningful parameters (i.e., strictly positive), so $\mathcal{O}$ is full rank. Thus, (A, C) is observable, and hence detectable. One can show that [19], the pair $(A, Q^{1/2})$ is stabilizable if

$$rank[A - \lambda I,\ Q^{1/2}] = 3 \ \ \forall \lambda \in \mathbb{C} : \lambda \geq 1 \quad (12)$$

Since for our system the matrix Q is diagonal, with positive entries on the diagonal, $Q^{1/2}$ is full rank, regardless of the what A and λ are. Therefore the pair $(A, Q^{1/2})$ is stabilizable.

Since (A, C) is observable and $(A, Q^{1/2})$ is stabilizable for any choice of the parameter set, $p = (C_z, C_w, R_z, R_w, A_e)$, that is physically meaningful (meaning all parameters are strictly positive), convergence and stability of the filter is guaranteed; no special care has to be take in choosing the parameter values.

4. Evaluation with simulation data

We now present results from applying the proposed method (SPDI Algorithm) to data generated from a simulation model. The simulation model itself is of form (1)-(2), so that there is an unambiguous ground truth that the results of the algorithm can be compared to. That is, the input q_{dist} for (1) will be compared to the estimated augmented state in (4). The parameters and the occupant-induced load are chosen somewhat arbitrarily, but ensuring that the model prediction is not too far from observations collected in a building that will be described in Section 5. We use a sampling period $t_s = 1/12$ hours (5 minutes).

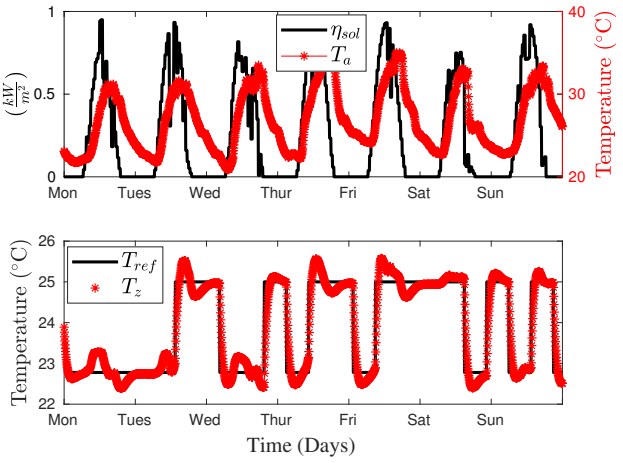

Figure 2: Some of the input data used in simulation evaluation: η_{sol}, T_a for all four scenarios (2013/08/26 to 2013/09/01), and T_z with T_{ref} for CL-NPW.

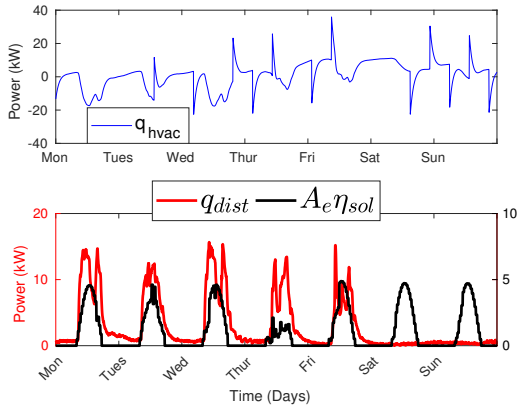

Figure 3: Heat gains q_{hvac}, $A_e\eta_{\text{sol}}$, and q_{dist} for the CL-NPW scenario in simulation evaluation. The heat gain q_{hvac} is a measured signal and is an input to the estimation method, while the other two are not known to the algorithm (though η_{sol} is).

Four scenarios are tested in simulation: (i): Open-loop simulation with a piecewise constant disturbance (OL-PW), (ii) Open-loop simulation with non piecewise constant disturbance (OL-NPW), (iii): Closed-loop simulation with a piecewise constant disturbance (CL-PW), and (iv): Closed-loop simulation with non piecewise constant disturbance (CL-NPW). For all four scenarios, the same input data sequences for ambient temperature, T_a, collected from `weatherunderground.com` and solar irradiance, η_{sol}, collected from the National Solar Radiation Database (NSRDB: `nsrdb.nrel.gov`), for Gainesville, FL, are used. In both the closed-loop scenarios, a PI controller is used to control the HVAC power q_{hvac} so that the room temperature tracks a set-point T_{ref}. The gains for this controller are elected to be 12 (proportional) and 4 (integral). To have exciting input to aid in identification, the reference room temperature trajectory is generated using a pseudo-random binary sequence (PRBS) [16, 10]. To ensure that occupant comfort is not compromised, the reference is constrained to lie within 22.78 °C and 25°C.

The inputs T_a and η_{sol}, that are used in all four scenarios, are shown in Figure 2(top). The output T_z for the CL-NPW scenario, along with the temperature set-point (T_{ref}), is shown in Figure 2 (bottom). A comparison of the various heat gains affecting the system, for the CL-NPW scenario, are shown in Figure 3. Note that the unmeasured disturbance q_{dist} is comparable to the magnitude of the cooling provided q_{hvac}. The presence of the large unmeasured disturbance means it cannot be ignored in estimating RC network parameters [10].

All simulation evaluations are done with two weeks of data, as explained in Section 3.1. All the plots in the paper show only the first week to reduce clutter.

4.1. Parameters

The parameters identified by the proposed method are shown in Table 1. In all scenarios, the parameters C_z, C_w, and A_e are identified quite accurately (maximum error relative to the true value of 1.6 %, 9.7 % and 6.0 % for C_z, C_w and A_e, respectively), but there are larger errors in the estimate of R_z and R_w in some scenarios: 35.7 % (R_z) and 20.0 % (R_w) in the worst cases. This can be explained from the structure of the model: the contribution of the term involving R_z and partly R_w in (2) may not be identifiable since this unknown term affects the wall state, which is furthermore not measured directly. Contrary to this, the parameter values C_z, and A_e affect a state equation that corresponds to the measured signal T_z, and they are estimated quite accurately in all four scenarios.

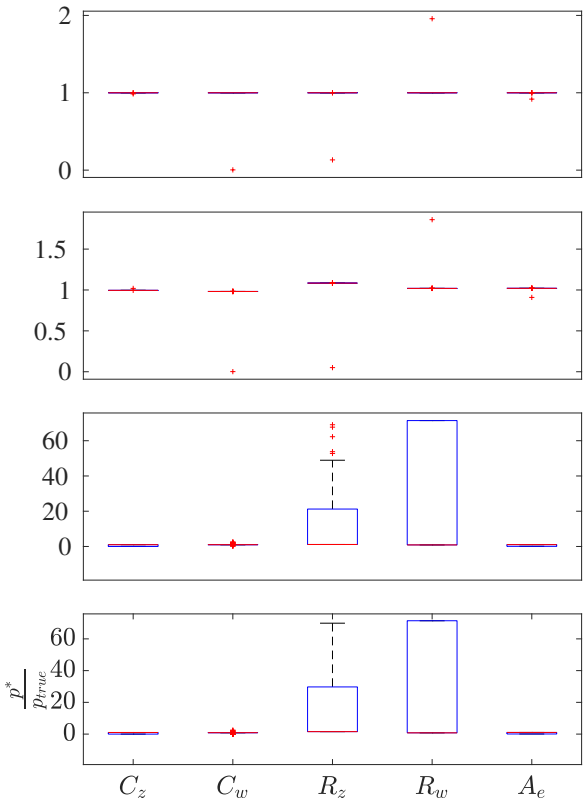

Figure 4: Simulation evaluation (all scenarios): Box-plot of optimal solutions of parameters normalized by true values for various initial guesses (from top: OL-PW, OL-NPW, CL-PW, CL-NPW).

Impact of non-convexity. Recall that the identified parameters are modes of the solutions to an optimization problem solved with n_{ig} (=50) distinct initial guesses. This was done to guard against getting wildly different estimates depending on the choice of initial guess to the underlying optimization problem, which is possible in principle since it is non-convex. To show the spread of these optima, box-plots of the optimal solutions are presented in Figure 4. From these plots it can be observed that all of the parameters from the open-loop scenario have little spread,

Table 1: Parameter Estimation results, $\hat{p}$.

Parameter	True Value	Estimates				units
		OL-PW	OL-NPW	CL-PW	CL-NPW	
C_z	9.40	9.40	9.35	9.55	9.54	kWh/K
C_w	20.00	20.00	19.64	20.40	18.06	kWh/K
R_z	0.70	0.70	0.76	0.79	1.09	K/kW
R_w	0.70	0.70	0.71	0.62	0.56	K/kW
A_e	7.00	7.00	7.14	7.42	7.27	m^2

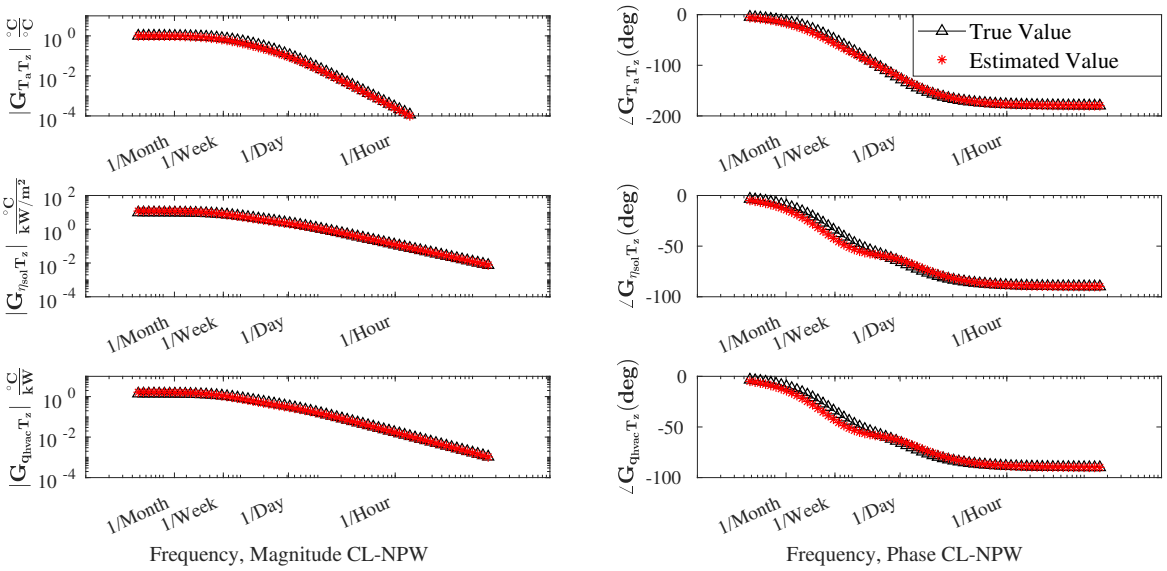

Figure 5: Simulation evaluation: Bode magnitude and phase plot comparisons for CL-NPW scenario.

however the closed-loop scenarios have considerable spread (particularly the two resistances). The non-convexity of the problem affects the closed-loop scenarios greater than the open-loop scenarios. This is expected as the system operating in closed-loop can be difficult to identify.

Even though many of the methods proposed in the literature to estimate R,C parameters require solving a non-convex optimization problem, the issue of non-unique solutions have not been treated in a principled manner.

4.2. Frequency response

Frequency response of a system refers to the gain and phase shift between the input and the output for sinusoidal inputs, as a function of frequency of the sinusoid. For prediction accuracy of the model, its frequency response is more important than individual parameters. Figure 5 shows the Bode plots, which are graphical representation of the frequency response, of the true and identified models. Due to space limitations, we only show the Bode plots for the CL-NPW scenario since that is the most relevant for practical applications.

Among the frequency responses, from inputs q_{hvac}, T_a, and η_{sol} to output T_z, the maximum absolute error between the identified and true response, observed over all frequencies, is found to be associated with the input T_a. This maximum error is 0.3071, or $\approx 31\%$ of the true value:

$$\max_{\omega} \frac{|\hat{G}_{\eta_{sol}T_z}(j\omega) - G_{\eta_{sol}T_z}(j\omega)|}{|G_{\eta_{sol}T_z}(j\omega)|} = 0.3071. \tag{13}$$

and occurs at the frequency of 1/50.3 hours. That the maximum error occurs for the input T_a is not surprising as the parameter with the second largest error, R_z, directly affects the corresponding transfer function.

4.3. Occupant-induced load

The estimated disturbance, $\hat{q}_{dist}$, for all four scenarios are shown in Figure 6. The estimated disturbance is quite accurate in almost all scenarios. It is especially surprising that even when the true disturbance is not piecewise constant, the estimated disturbance is able to capture the main trend of the disturbance well. The estimates are slightly better in the open-loop scenarios than in the closed-loop scenarios, which is not surprising since system identification with closed-loop data is more challenging.

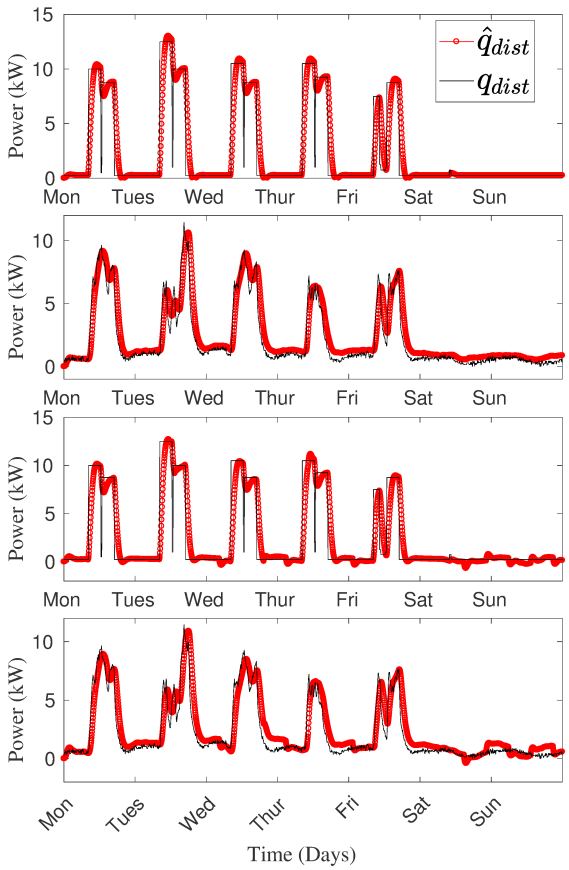

Figure 6: Simulation evaluation: estimated q_{dist} for all four scenarios (top to bottom: OL-PW, OL-NPW, CL-PW, CL-NPW).

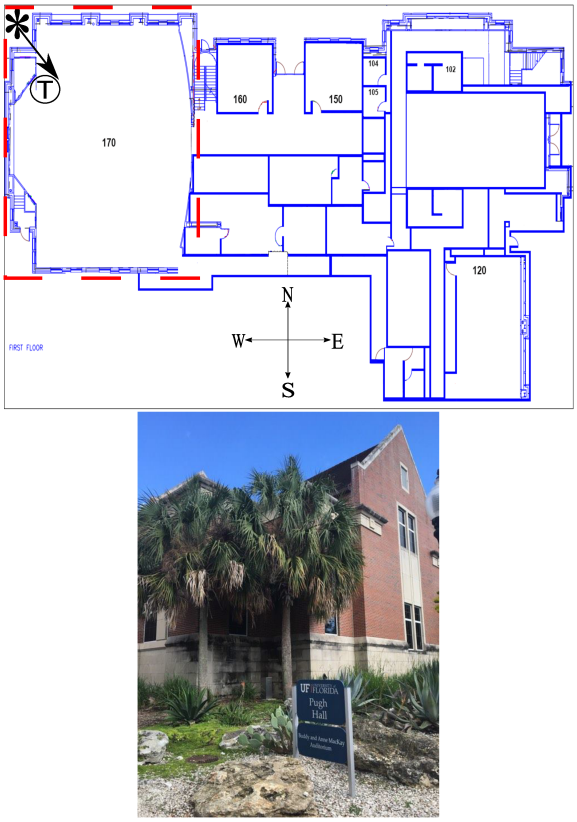

Figure 7: Pugh hall floor plan (top) and photograph (bottom), with the zone that the SPDI algorithm is applied to is shown enclosed in dashed lines. The "*" denotes the location from where the photograph of the building was taken (show at the bottom) and the arrow denotes direction of the camera. The Ⓣ is the location of the temperature sensor that is used for T_z.

5. Evaluation with building data

5.1. Building Description

The proposed method has been applied to a zone in a building in the University of Florida campus. In particular, data from a large auditorium ($\sim$ 540 m^2) that is serviced by a dedicated air handling unit is used. This zone, which serves as a building for this study, is shown within the red dashed line in Figure 7. The solar irradiance η_{sol} data has already been described in Section 4. The input q_{hvac} (see top of Figure 8) was computed from measurements of supply air flow rate m_{sa} and temperature T_{sa} (measured at the supply duct), and of return temperature T_{ret} (measured at the return duct): $q_{\text{hvac}}(t) = m_{sa}(t)C_{pa}(T_{sa}(t) - T_{ret}(t))$, where C_{pa} is the specific heat capacity of air. Note that latent heat is ignored; so the calculation is accurate only when there is little difference in the supply and return duct humidities. The zone temperature, T_z, is measured at a thermostat in the zone, whose location is denoted by the circled T in Figure 7. The data used here was collected during normal operation where the HVAC system is under closed-loop control, no special excitation test was done to aid the estimation algorithm.

Table 2: Parameter estimates from closed-loop data collected from Pugh Hall.

Parameter	Estimate	units
C_z	8.70	kWh/K
C_w	19.70	kWh/K
R_z	0.60	K/kW
R_w	0.55	K/kW
A_e	8.12	m^2

5.2. Parameters

The estimated parameters are shown in Table 2. Though not much can be said about the accuracy of the parameter estimates due to lack of ground truth, the relative magnitude of the parameters appear reasonable. The purpose of the wall state is to model the slower time-scale thermal dynamics while the room state is to model the faster time-scale, and the estimated parameters are consistent with this: estimate of C_w is about 2 times larger than C_z. In addition, the estimate of C_z is close to the value used in simulations, which was arrived at by manual calibration with data from this building. The two resistance values, R_z and R_w, are estimated to be close to each other.

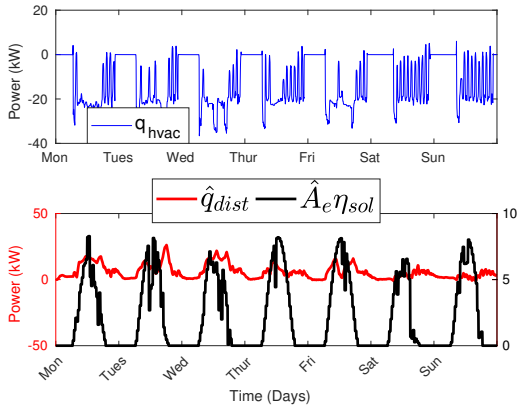

Figure 8: Evaluation with building data: Heat gains q_{hvac}, $\hat{A}_e\eta_{sol}$, and $\hat{q}_{dist}$ estimate for 2013/08/26 to 2013/09/01.

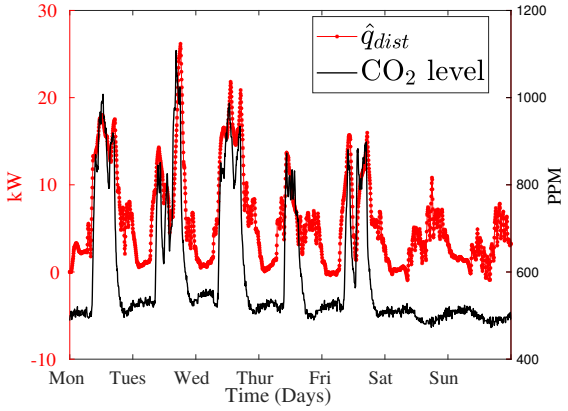

Figure 9: Evaluation with building data: $\hat{q}_{dist}$ vs. measured CO_2 concentration, for the week of 2013/08/26 to 2013/09/01.

5.3. Occupant-induced load

The q_{dist} estimation results are compared against CO_2 (measured at the return duct) in Figure 9 and the estimated solar heat gain $\hat{A}_e\eta_{\text{sol}}$ in Figure 8 (bottom). The estimated disturbance correlates strongly with the CO_2 concentration measurements during the first weekdays (first five days), with correlation coefficient being 0.83. Remarkably, during the week days the estimated occupant-induced load captures the trend of the CO_2 signal well. However, the flat CO_2 measurement during the weekend indicates the building is empty and thus the true occupant-induced disturbance is likely to be nearly constant or zero, while the estimated disturbance is not. The difference could be due to unmodeled dynamics (such as latent heat, ground coupling etc.). It could also be due to the Kalman filter, which has a tendency to use the third state q_{dist} as a "garbage collector". Recall that a large variance is assigned to process noise corresponding to the the third state, q_{dist}, in the Kalman filter equations. A large variance is needed for the model to be able to track the time-variations in the true occupant-induced disturbance that deviates from the assumed piecewise-constant profile. That inherently limits the model's ability to accurately estimate the constant q_{dist} when it is indeed constant, such as during weekends.

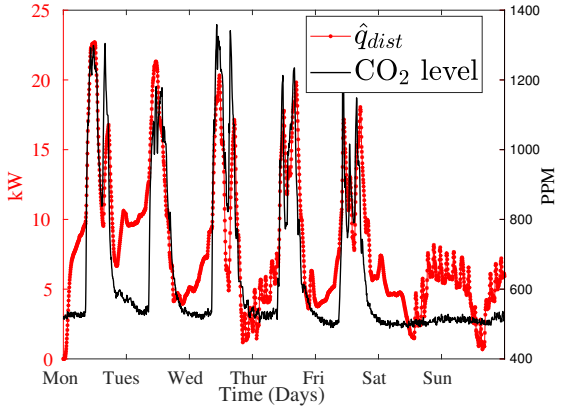

Figure 10: Evaluation with building data (winter): $\hat{q}_{dist}$ vs. measured CO_2 concentration, for the week of 2013/01/14 to 2013/01/20.

For additional assessment of the quality of disturbance estimates, the parameter values in Table 2 were used with (4) to estimate the internal disturbance, q_{dist} for another week of data. Since the parameters are known, the problem reduces to one of state estimation. The load estimates are shown in Figure 10. The disturbance estimate is again strongly correlated with the CO_2 signal during the weekdays. The correlation coefficient is now 0.83, the same as in the summer data, even though this data set is from the winter.

5.4. Root mean square error

Since no ground truth is available to test the plausibility of the model fitted to building data, the root mean squared error (RMSE) of the residuals of the fit (r_k in (6)) is computed: $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} r_k^2}$. For the first week of data, the RMSE is 0.23°C.

6. Conclusion

We present a method to simultaneously estimate (i) the parameters of a dynamic model of a building's temperature including resistance, capacitances and effective solar area, and (ii) unmeasured disturbance (heat gain), from input-output data. The method is based on the assumption that the main component of the unmeasured disturbance, the occupant-induced load, is piecewise constant, especially in commercial buildings. Extensive simulation tests indicate that the method works well even when the true occupant-induced load is not piecewise constant, and with both open-loop and closed-loop data, making it highly applicable in practice. When applied to data collected from a real building, the method appears to work well in the following sense. One, the values of the capacitance's represent appropriate magnitudes when it comes to modeling the slow and fast time-scales of temperature variation.

Two, the estimated heat-gain is strongly correlated with measured CO_2 concentration. The true occupant-induced load should be strongly correlated to the CO_2 concentration.

There are several avenues of future work. The model structure used here is linear, which comes from lumping the cooling injected by the HVAC system, into one signal [10]. Extending the method to a model in which the supply air temperature and supply air flow rate are inputs, in which case the model becomes bilinear, will make it more useful for use in model-based control [21]. Finally, the model structure used here lumps an entire building into a single zone. Extending the method to multi-zone buildings is another avenue for future work.

References

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

[2] F. B. Morris, J. E. Braun, S. J. Treado, Experimental and simulated performance of optimal control of building thermal storage, ASHRAE transactions 100 (1) (1994) 402–414.

[3] H. Hao, T. Middelkoop, P. Barooah, S. Meyn, How demand response from commercial buildings will provide the regulation needs of the grid, in: 50th Annual Allerton Conference on Communication, Control and Computing, 2012, invited paper.

[4] J. Penman, Second order system identification in the thermal response of a working school, Building and Environment 25 (2) (1990) 105 – 110.

[5] H. Madsen, J. Holst, Estimation of continuous-time models for the heat dynamics of a building, Energy and Buildings 22 (1) (1995) 67 – 79.

[6] 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.

[7] S. Wang, X. Xu, Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm, Energy Conversion and Management 47 (1314) (2006) 1927 – 1941.

[8] Y. Lin, T. Middelkoop, P. Barooah, Issues in identification of control-oriented thermal models of zones in multi-zone buildings, in: IEEE Conference on Decision and Control, 2012, pp. 6932 – 6937. doi:10.1109/CDC.2012.6425958.

[9] X. Li, J. Wen, Review of building energy modeling for control and operation, Renewable and Sustainable Energy Review 37 (2014) 517537.

[10] 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. doi:http://dx.doi.org/10.1016/j.buildenv.2016.07.007.

[11] Q. Hu, F. Oldewurtel, M. Balandat, E. Vrettos, D. Zhou, C. J. Tomlin, Building model identification during regular operation-empirical results and challenges, in: Proceedings of the American Control Conference, IEEE, 2016, pp. 605–610.

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

[13] S. F. Fux, A. Ashouri, M. J. Benz, L. Guzzella, EKF based self-adaptive thermal model for a passive house, Energy and Buildings 68, Part C (2014) 811 – 817.

[14] A. Leavey, Y. Fu, M. Sha, A. Kutta, C. Lu, W. Wang, B. Drake, Y. Chen, P. Biswas, Air quality metrics and wireless technology to maximize the energy efficiency of {HVAC} in a working auditorium, Building and Environment 85 (2015) 287 – 297.

[15] J. Chen, C. Ahn, Assessing occupants energy load variation through existing wireless network infrastructure in commercial and educational buildings, Energy and Buildings 82 (2014) 540 – 549. doi:https://doi.org/10.1016/j.enbuild.2014.07.053.

[16] L. Ljung, System Identification: Theory for the User, 2nd Edition, Prentice Hall, 1999.

[17] MATLAB fmincon, https://www.mathworks.com/help/optim/ug/fmincon.html, accessed: 2017-05-04.

[18] M. D. McKay, W. J. Conover, R. J. Beckman, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics 21 (1979) 239.

[19] J. Hespanha, Linear systems theory, Princeton University Press, 2009.

[20] B. Anderson, J. Moore, Optimal Filtering, Prentice-Hall, Englewood Cliffs, NJ, 1979.

[21] S. Goyal, P. Barooah, A method for model-reduction of nonlinear building thermal dynamics of multi-zone buildings, Energy and Buildings 47 (2012) 332–340.