Identification of multi-zone building thermal interaction model from data

Siddharth Goyal, Chenda Liao and Prabir Barooah

***Abstract*— Constructing a model of thermal dynamics of a multi-zone building requires modeling heat conduction through walls as well as convection due to air-flows among the zones. Reduced order models of conduction in terms of RC-networks are well established, while currently the only way to model convection is through CFD (Computational Fluid Dynamics). This limits convection models to a single zone or a small number of zones in a building. In this paper we present a novel method of identifying a reduced order thermal model of a multi-zone building from measured space temperature data. The method consists of first identifying the underlying network structure, in particular, the paths of convective interaction among zones, which corresponds to edges of a building graph. Convective interaction among a pair of zones is modeled as a RC network, in a manner analogous to conduction models. The second step of the proposed method involves estimating the parameters of the RC network model for the convection edges. The identified convection edges, along with the associated R and C values, are used to augment a thermal dynamics model of a building that is originally constructed to model only conduction. Predictions by the augmented model and the conduction-only model are compared with space temperatures measured in a multi-zone building in the University of Florida campus. The identified model is seen to predict the temperatures more accurately than a conduction-only model.**

I. INTRODUCTION

Buildings are one of the primary consumers of energy worldwide. Inefficiencies in building technologies, particularly in operating the HVAC (heating, ventilation and air conditioning) systems cause a significant fraction of the energy consumed by buildings to be wasted. As a result, there is a growing interest in developing techniques that can compute optimal building control signals to minimize building-wide energy consumption, such as MPC (model predictive control) [1], [2]. These control techniques require a model that adequately captures the relevant dynamics of a building, especially of the thermal dynamics that relate the control signals to the space temperatures, i.e., average temperatures of the zones of the building. In addition, the models should have a small state space dimension. Otherwise, computing the control signals becomes computationally expensive and in some cases, infeasible. Quite a few building energy simulation programs are available, such as EnergyPlus [3], TRNSYS [4], and DOE-2 [5]. Although these programs are useful for load calculations, equipment sizing, and predicting energy use of a building over long time intervals, their utility is limited as tools to model or simulate the dynamics of the thermal processes inside a building that can be used by a control system [6].

Constructing building thermal dynamics is a challenging task since it requires modeling heat exchange through convection, conduction and radiation among all the rooms. The thermal dynamics in a multi-zone building can be thought of as an interconnected system of many subsystems. Each subsystem corresponds to a zone, and the interconnections correspond to dynamic interactions between pairs of zones, which may occur due to conduction or convection. A first-principles based model constructed from energy and mass balance equations will lead to a highly complex model. Currently, modeling convection requires CFD simulations, which limits its application to one zone or a small number of zones [7]. In contrast, modeling heat exchange between zones due to conduction is quite feasible; substantial literature exists on modeling conduction using RC (resistor capacitor) networks. However, little work has been done on constructing reduced order models of convection. The key challenge is therefore to construct a reduced order model of convective interactions among the zones of a building. Due to the complexity of the underlying physics, a data-driven approach that identifies these interactions from observed behavior is more likely to succeed than a physics-based one. Even with a data driven approach, there are two main challenges. The first is to determine which pairs of zones have significant convective interaction in a multi-zone building. The second is to develop a reduced order model for convective heat transfer between a pair of rooms. If these two challenges are overcome, a model of the whole building can be constructed as a network of elements, each element being a reduced order model of either convection or conduction between a pair of zones.

A network model of thermal dynamics of a multi-zone building will have nodes corresponding to the temperatures in zones and edges corresponding to reduced order models of dynamic interaction between the variables connected by the edge. In this paper, we address the problem of identifying the network model of thermal dynamics from measured zone temperatures and input signals. The edges in the network that correspond to the conduction are straightforward to determine from the building's geometry. The edges that correspond to convection are far more challenging to identify. We borrow ideas from machine learning, in particular, concentration graph models, to determine these edges, which are based on computing conditional dependencies among zone temperatures. Reduced order models of conduction

This work has been supported by the National Science Foundation by Grants CNS-0931885 and ECCS-0955023. The authors are with the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, Florida, USA. {siddgoya,cdliao,pbarooah}@ufl.edu

in terms of RC-circuits are well-established [8], [9], [10]. Motivated by RC-circuit models of conduction, we model convection between two zones as a RC circuit as well. The resistance and capacitance values of convection edges are obtained minimizing a prediction error. The proposed method is applied to a section of a building in the University of Florida campus. Comparison of the model's prediction with measured data shows that the identified model predicts the temperatures more accurately than a model that only takes conduction into account.

The rest of the paper is organized as follows. Section II formulates the problem precisely. Section III describes the proposed method. Performance of the method when applied to a section of a building in the University of Florida campus is discussed in Section IV. The paper ends with a discussion on future research directions in Section V.

II. Thermal network model and problem formulation

A commercial building is usually divided into a number of "zones"; where a zone is either a number of rooms or a single large room. Conditioned air is supplied to the "terminal boxes" of respective zones. The flow rate and temperature of this air are varied at the terminal boxes through dampers and reheat coils, before being supplied to the zone, to maintain the zone temperature at a desired value. For modeling purposes, we assume that the air in the zone is well mixed. The zone temperature T_i and humidity ratio W_i of i^{th} zone are state variables of a thermal dynamics model.

The *network* nature of the thermal dynamics model comes from the fact the states (temperature and humidity) of a zone are affected by the states of nearby zones due to conduction, convection, and radiation. In this paper, heat exchange among zones due to radiation is neglected; which is also a common practice [11]. We use the commonly used 3R2C reduced order model of conduction between two spaces separated by a solid surface [10], [8]. However, modeling convection is quite challenging. Typically, convection is analyzed through CFD simulations, since the governing equations are a set of coupled partial differential equations [7], [12], [13]. However, this approach is limited to a single zone or a very small number of zones due to computational complexity. To the best of our knowledge, no work has been done on constructing *reduced-order dynamic models of convection in multi-zone buildings*. As we are interested in lumped parameter, or reduced-order models, in this paper, we model convective interaction between two zones by a RC circuit as well.

Thermal interaction among multiple zones in a building can now be described in terms of a undirected graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ with node set $\mathbf{V} = \{1, \ldots, n\}$ and edge set $\mathbf{E} \subset \mathbf{V} \times \mathbf{V}$. Each node in the set $\mathbf{V}$ corresponds to a variable, e.g. temperature in a room. If the variables corresponding to nodes u and v directly affect each other, then we say that there is an edge between u and v: $(u, v) \in \mathbf{E}$. In general, a node can correspond to variables such as temperature, humidity ratio, etc. In this study, all the nodes will correspond to temperatures at certain

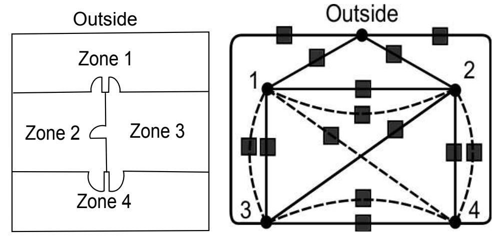

Fig. 1. A schematic illustration of a thermal network model, with edges representing either convective or conductive interaction by solid and dashed lines respectively. Each square solid box is a RC-circuit.

locations inside or outside the building. Hence, the number of nodes, n, in the network model is equal to the number of zones in the building in this paper. An edge is essentially a path for dynamic interaction between two variables that are relevant to the thermal dynamics. For example, if a pair of nodes u and v correspond to two adjacent rooms separated by a wall so that there is conductive heat transfer between them, or if there is air flow between u and v, then this is represented by an edge between the nodes u and v. The edges that arise due to conductive heat exchange are called *conduction edges*, while those that arise out of convective heat exchange between two zones (or between a zone and the outside, though rare) are called *convection edges*. A schematic illustration of a thermal network model of a four-zone building is shown in Figure 1, where solid and dashed line represent the conduction and convection edges respectively. Each edge has a RC-circuit associated with it, which models the dynamic interactions between the variables that are associated with the corresponding pair of nodes. Since we model both conduction and convection using a 3R2C model, each edge has a LTI dynamic system of state-space dimension 2 associated with it. Each node also has a state associated with it: the temperature of the zone that corresponds to the node. We call the zone temperatures the *node variables*. In summary, a thermal network model of a multi-zone building is a graph along with weights on the edges, where each "weight" is an LTI system that corresponds to the differential equations of a RC network. We refer the reader to [14] that describes the model in detail, and to [15], [16] for model reduction techniques developed for such network models.

A. The identification problem

Identification of a thermal network model consists of (i) identifying a minimal set of edges between node pairs, and (ii) estimating the parameters for each edge (i.e., the R and C values for the RC-circuits associated with the edges) required to explain the observed behavior of the node variables. We call the first the *structure identification* problem and the second the *parameter estimation* problem. Identifying the conduction edges and the R and C values associated with it is straightforward [10]. However, identifying the convection edges and determining their R and C values is far more challenging, and forms the crux of the problem we address in this paper.

III. Proposed Identification Method

As described in the previous section, identification of a thermal network model from measured input-output data can be thought of in terms of two sub-problems: (i) structure identification and (ii) parameter estimation. We now describe our approach to solve these two sub-problems.

A. Network structure identification

The proposed method for structure identification relies on identifying conditional independence between random variables. Recall that two random variables (r.v.) X and Y are called conditionally independent given a r.v. Z, if they are independent given the σ-algebra generated by Z. For instance, the temperatures in two buildings that are physically separate but are located in the same neighborhood are conditionally independent given the outside air temperature, assuming that apart from the outside temperature there is no common input that affects both the buildings. Similar arguments can be made for temperatures of zones in a multi-zone building. If T_i and T_j are node variables (i.e., zone temperatures) that are conditionally independent given the rest of the node variables, that would mean that there is no direct thermal interaction between these two zones. So determination of conditional independence between pairs of node variables (given all the rest) provides us a way to determine which pairs of node variables should not have edges between them in the thermal network, and vice versa.

To determine pairwise conditional independence, we use the idea of the so-called *concentration graph* ($\mathbf{G}_c$) model from machine learning [17]. Given n random variables $X_1,\ldots,X_n$ that are jointly Gaussian, the inverse-covariance matrix $P := \Sigma^{-1}$, where $\Sigma = Cov(\mathbf{X},\mathbf{X})$ and $\mathbf{X} = [X_1,\ldots,X_n]^T$, provides information on conditional dependencies. In particular, two variables X_i and X_j are conditionally independent given the rest if $P_{i,j} = 0$. For minimum model complexity, it is also desired that the estimated concentration graph have as few edges as possible, i.e., the estimated P should be as sparse as possible. The need of sparsity becomes more important as the number of variables becomes large.

From the maximum likelihood estimator (S) of the covariance Σ, the estimated concentration matrix can be obtained as $\hat{P} = S^{-1}$, which is not sparse in general. The method in [17], which we will utilize, leads to an estimated graph model $\hat{\mathbf{G}}_c$. We first identify the concentration graph model of the node variables $T_i,\ i = 1,\ldots,n$ from time-series data using the method of [17]. In fact, if measurements are collected at m-minute intervals so that we have $K = \frac{24\times 60}{m}$ discrete time indices for 24 hours, then we identify K distinct concentration graph models, $\hat{\mathbf{G}}_c(k)$, where $k = 1,\ldots,K$ is the time index. We treat each day's data as an independent realization of the underlying stochastic processes. Therefore, given N days of temperature data, the estimate of $P(k)$ for each k is based on N samples. If $P_{ij}(k) \neq 0$ where T_i and T_j are node variables of i and j in the thermal network, then we conclude that these two node variables directly affect each other at time k, and therefore (i,j) is a potential edge. Among these potential edges, the edges that correspond to the conduction, which are already known from building geometry, are first chosen. The convection edges are chosen next, as follows. Let $\bar{G}(k) = (\mathbf{V},\bar{\mathbf{E}}(k))$, where $\bar{\mathbf{E}}(k)$ is $\mathbf{E}(k)$ with all the conduction edges removed. For a fixed node i, we determine the score of all other nodes j ($j = 1,\ldots,n$) as:

$$s_j^{(i)} := \frac{\#\text{ of times } (i,j) \text{ appear in } \cup_k \bar{\mathbf{E}}(k)}{K}.$$

If node ℓ has the highest score: $\ell = \arg\max_j s_j^{(i)}$, then (i,j) is chosen as the convection edge for i. This process is repeated for each i. Note that only one convection edge is chosen for each node at each stage, though at the end of the process a node may have multiple convection edges incident on it. This is done in order to restrict the number of convection edges, and thereby obtain a model that is of minimal complexity.

B. Parameter estimation

Once the edges of the network model are identified, the parameters corresponding to each edge (3 R values and 2 C values) have to be estimated. We choose the zone capacitance to be of the same order of magnitude of internal wall capacitance, and proportional to the volume of corresponding zone. Since we model each convection edge as a 3R2C circuit as well, three R values and two C values need to be estimated for each such edge. For the sake of simplicity, we first restrict all the 3 resistors to have the same resistance and both the capacitors to have the same capacitance. From now on, we use subscripts d and v to refer to conduction and convection, respectively. Thus, we only need to estimate only two parameters, R_v and C_v, for each the convection edge.

The time constant τ of a 3R2C circuit with all three resistances equal to R and both the capacitances equal to C is proportional to RC, so that $\tau = \alpha RC$. So, if such a circuit is used to model heat transfer (conduction or convection), we have

$$\tau_\mathrm{d} = \alpha R_\mathrm{d} C_\mathrm{d}, \qquad \tau_\mathrm{v} = \alpha R_\mathrm{v} C_\mathrm{v} \tag{1}$$

where $\tau_{(\cdot)}$ refers to time constant, α is the constant of proportionality, and the subscripts d and v refer to conduction and convection, respectively. The time constant of a model of convective heat transfer between two zones should increase with the physical distance between them. The larger the distance, the longer it will take to transfer heat, since convection requires physical mass exchange. We incorporate this effect by making the resistance of the RC-circuit model for convection proportional to d, the physical distance between the two zones along the most likely path of air exchange (which is usually a hallway): $R_\mathrm{v} = \bar{R}_\mathrm{v} d$, where the proportionality constant $\bar{R}_\mathrm{v}$ can be thought of as resistance per unit distance. We assume that the capacitance C_v is not affected by the distance between the zones. Now, (1) can be rewritten as

$$\tau_\mathrm{v} = \alpha d \bar{R}_\mathrm{v} C_\mathrm{v} \tag{2}$$

To fully specify a RC network model of a convective edge in the network model, the parameters α, $\bar{R}_\mathrm{v}$, and C_v are needed; d is known from building geometry. The value of α can be computed from (1) since the parameters R_d and C_d are known for a surface of specific material and geometry, and $\tau_{\rm d}$ can be determined as the absolute value of the inverse of the least stable eigenvalue of the LTI system corresponding to a 3R2C network model for that surface. The remaining parameters $\bar{R}_{\rm v}$, $C_{\rm v}$ are chosen by searching for values that minimize the prediction cost J:

$$J := \sum_{k=1}^{K_T} \sum_{i \in \mathbf{Z}} (\hat{T}_i(k) - T_i(k))^2 \tag{3}$$

where $T_i(k)$ and $\hat{T}_i(k)$ are the measured and predicted temperatures respectively of i^{th} room at time index k, K_T is the total number of time steps over which measurements and model predictions are obtained, and $\mathbf{Z} \subset \mathbf{V}$ is a subset of the zones in the building. Temperature predictions are obtained from simulating the model described in [14].

Since nothing is known about the structure of this optimization problem, in this study we obtain the parameters $\bar{R}_{\rm v}$, $C_{\rm v}$ by an exhaustive search. Knowledge of approximate values of the time constant of convective heat exchange is used to constrain the search space. The change in a zone's temperature due to convection is much faster than that due to conduction. In a CFD-based study of convection in multi-zone buildings reported in [7], the time constant due to convection across two locations 2.5 m apart is seen to be between 10 seconds and 50 seconds depending on the locations. We call $\frac{1}{d}\tau_{\rm v}$ as the *time constant per unit distance*, and from the results reported in [7], we impose the condition that $t_{\min} \leq \frac{1}{d}\tau_{\rm v} \leq t_{\max}$, where $t_{\min}$ and $t_{\max}$ are the lower and upper bounds on the time constant per unit distance, for convection. It follows from (2) that $\frac{1}{\alpha}t_{\min} \leq \bar{R}_{\rm v}C_{\rm v} \leq \frac{1}{\alpha}t_{\max}$. For the choice of values $t_{\min} = 4$, $t_{\max} = 20$ (which follow from the results in [7]), we obtain $0.4 \leq \bar{R}_{\rm v}C_{\rm v} \leq 2$. The value of α used in this calculation is 9.86, which is computed as described earlier. Hence, in searching for the values of $\bar{R}_{\rm v}, C_{\rm v}$ that minimize the prediction cost J, the search was limited to the range $\bar{R}_{\rm v}C_{\rm v} \in [0.4 \quad 2]$.

IV. Results

The method described above is applied to identify the thermal network mode of a section of a building (Pugh Hall) located at the University of Florida campus, Gainesville, FL. The section of the building chosen is a part of the second floor of Pugh Hall; its layout is shown in Figure 2. The node variables are the temperatures of the 7 zones that are denoted as 200, 210, 230, 245, 248, 249 and "Hallway" in Figure 2. The reason for choosing this section of the building as a test-case is the availability of the time-series data for the boundary nodes, i.e., of the zones 200, 210, 230, 248 and corridor as well as outside. Measurements of zone temperatures, supply air temperatures and flow rates are obtained from the Siemens Insight© BAS (Building Automation System), at 5 minutes intervals for 26 days starting from January 21, 2011. The outdoor temperature data is obtained from [18] at 60 minute intervals for the same period.

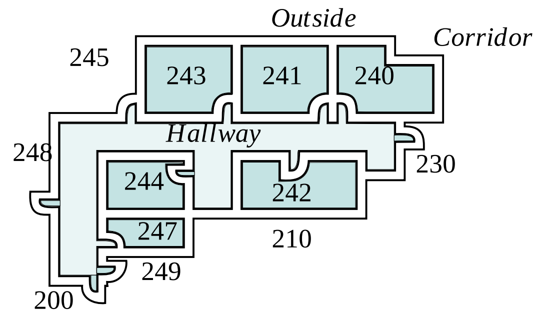

Fig. 2. A detailed schematic of the 7-zone section on the 2^{nd} floor of Pugh Hall, 40000 sq. ft. building located in the University of Florida campus.

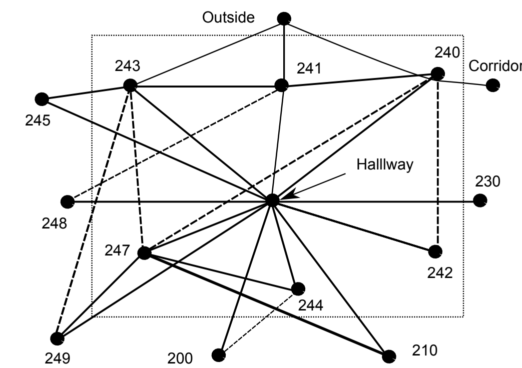

Fig. 3. Results of network structure identification method applied to find the convection edges for a block shown in Figure 2. The conduction and convection edges are shown as solid and dashed lines, respectively.

A. Model simulation details

To simulate the network model in MATLAB©, after it is identified, inputs and initial conditions are required. Each zone has sensors that measure supply air temperature, supply air flow rate, and temperature of the zone, but there are no sensors to measure humidity ratios. Thus, among the inputs to the model, the supply air temperature and flow rate, and temperatures of the boundary nodes are known, while the supply air humidity ratios are not. Hence we simulate the model assuming that humidity ratio of conditioned air supplied into each zone is the same and is constant over time. This is a reasonable assumption since air handling units deliver air at an approximately constant humidity ratio, and reheating at the terminal boxes does not affect humidity. The constant value of humidity ratio was chosen after on-site measurements as 0.0074. The initial zone temperatures in the model are chosen to be the same as the measured initial temperatures. The initial values of the internal states of each conduction edge are chosen as the average of the initial temperatures of the zones connecting them [14]. The section of the building used in the current study has only three rooms that share walls with the outside, and furthermore these walls are north facing, with no direct sunlight incident on them. Therefore, we assume that the solar load in all the zones are 0 at all times. Furthermore, all the rooms in the section are offices with a designed average occupancy of 2. Simulations of the identified model (which will be described next) indicated that the difference in the temperature prediction with two occupants per room and no occupants is less than $0.3^{\circ}F$. Hence, loads from occupants are set to 0 in all the simulations. Since there is no sensor to measure the hallway temperature, we assume that the initial hallway temperature as the average of the initial temperatures of the zones next to the hallway. It is assumed that the thermal resistance of the floor and the roof of 2^{nd} floor of the building is much larger than that of internal walls. The latter means that temperatures of 1^{st} and 3^{rd} floors of the building have little impact on the temperatures of the zones in the 2^{nd} floor.

B. Identification and verification

The available data for the building section shown in Figure 2 is separated into a calibration data set (data for January 21, 2011 through February 9, 2011) and a validation data set (rest of the data). Network structure is identified using the whole calibration data set and parameters are estimated using only the first 12 hours of the calibration data set. The resulting identified graph is shown in Figure 3, where the nodes lying outside the box shown in dashed line correspond to boundary nodes. In Figure 3, a solid line represents the conduction edge and a dashed line represents a convection edge. To estimate the values of $\bar{R}_{\text{v}}$ and C_{v} for the convection edges, the value of the prediction cost J is computed for various values of these parameters, by varying $\bar{R}_{\text{v}}$ between $10^{(-6)}$ and 0.01 and varying $\tau_0 := \bar{R}_{\text{v}} C_{\text{v}}$ between its allowable values (see Section III-B). The length of the time interval used in computing J is $k_T = 144$, which corresponds to first 12 hours of the calibration data set. Model predictions for a given set of parameter values are obtained from MATLAB©simulations of the model, as described in Section IV-A. Figure 4 shows the variation of cost functional J defined in (3) as a function of the resistance $\bar{R}_{\text{v}}$. The cost functional achieves a minimum at $\bar{R}_{\text{v}} = 3.36 \times 10^{-4}$, and the corresponding capacitance is $C_{\text{v}} = 1.19 \times 10^{3}$, which are therefore chosen as the estimated parameters. If we introduce the convection edges manually with with R and C values estimated above, the same convection edges are recovered back when the same method is applied.

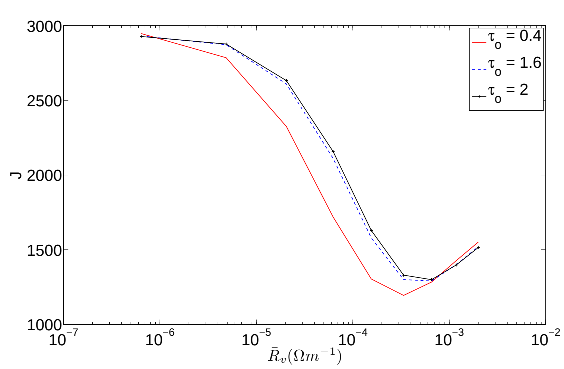

Fig. 4. Value of objective function J defined in (3) as a function of the convection edge resistance R, for a few values of the time constant.

C. Validation

Once the network model (structure and parameters) are determined as described above, the model is simulated in MATLAB©for a given set of inputs and the predicted outputs are compared with measured values. The inputs used for these simulations are obtained from measured data during midnight of February 10, 2011 to midnight of February 12, 2011, which is part of the validation data set. Figures 5 and 6 show the measured temperatures, the temperatures predicted by the identified network model, and those by a conduction-only model, of room 243 and 244. All time traces shown are for the 48 hour time period mentioned above. It is clear that the temperature predictions by the identified network model that includes convection effects are substantially closer to the measured values than those by the conduction-only model. The maximum error between measured values and prediction of the identified network model is about $3^{0}F$, while the maximum error is about $6^{0}F$ if only conduction is taken into account. The predictions by the identified model for the other rooms (not shown due to space limitations) are also closer to the measured values than the predictions by the conduction-only model.

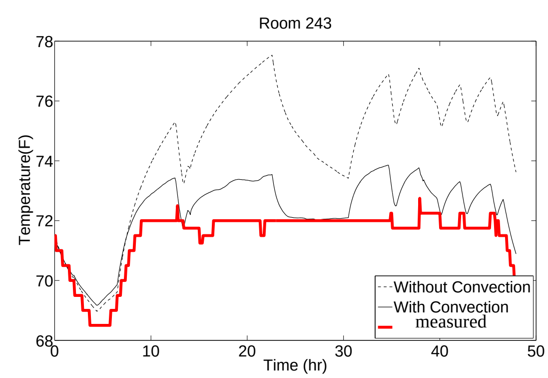

Fig. 5. Actual and predicted temperatures of room 243 with the identified model and with conduction-only model.

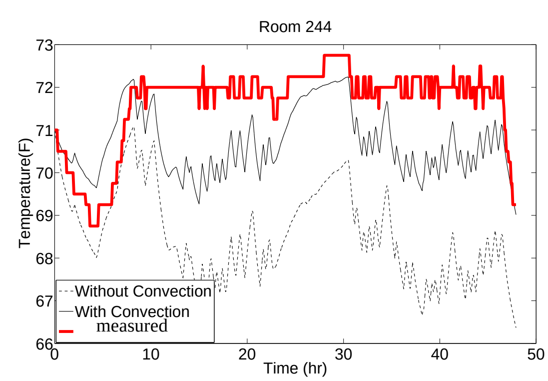

Fig. 6. Actual and predicted temperatures of room 244 with the identified model and with conduction-only model.

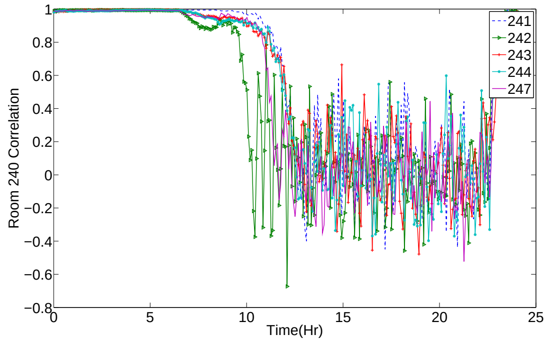

Fig. 7. Correlation coefficients between pairs for rooms (240, i), where i = 241, 242, 243, 247 over a 24 hour period (01/20/2011-02/15/2011). The estimates are computed from data collected over 26 days. This plot shows that marginal dependencies between zone temperatures do not help in unearthing thermal interactions among zones.

V. Discussion and future work

We proposed a method to identify the interconnection structure of the thermal dynamic model of a multi-zone building. The proposed method relies on estimating conditional independence between pairs of zone temperatures to estimate the convection edges. An additional contribution is a RC network based reduced order model of convection. The R and C values of the convection edges are estimated through an exhaustive search to minimize a prediction error cost. The identified model predicts the zone temperatures more accurately compared to a conduction-only model.

It should be noted that it is important to examine conditional dependencies rather than marginal dependencies. Pairwise correlation coefficients among the node temperatures reveal little about cause and effect. Figure 7 shows the estimated correlation coefficients among pairs of zones in the building in Figure 2, which indicates that all the node variables are highly correlated during the nighttime and have little correlation during daytime. This is an artifact of the way the building's HVAC system is operated. The zone temperatures are allowed to "float" at night when the building is unoccupied. As a result, all the room temperatures tend to either increase or decrease together depending the outside temperature, which makes them all highly correlated. During the daytime, on the other hand, all the room temperatures are maintained close to 72^{o}F, with small random fluctuations that arise due to occupants and other loads. As a result the correlations among the zone temperatures during daytime are close to 0. Hence, little information on the interconnection structure of the thermal network can be obtained from examining correlation coefficients.

Numerous avenues for improvement exist, we list a few. In general, it is quite possible that two rooms have both conductive and convective interaction between them, but the proposed method does not produce parallel edges. A strength of the method is that it does not require any forced-response experiments to reveal the interconnections. However, since the zones of a building are usually maintained at a constant temperature, there is little "persistency of excitation" in the measured signals. There might be a limit on how much of the interconnection structure can be unearthed with such closed-loop data. This needs to be explored in the future.

Acknowledgments

The authors gratefully acknowledge the help of Prof. Herbert Ingley for many useful discussions on HVAC systems, John Mocko, for his in getting high resolution weather data, and the Physical Plants Division of the University of Florida, in particular, Skip Rockwell, for help in collecting data from Pugh Hall's building automation system.

References

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

[2] D. Gyalistras and M. Gwerder, "Use of weather and occupancy forecasts for optimal building climate control (opticontrol): Two years progress report," Siemens Switzerland Ltd, Tech. Rep., 2010.

[3] D. Crawley, L. Lawrie, C. Pedersen, and F. Winkelmann, "EnergyPlus: Energy simulation program," *ASHRAE Journal*, vol. 42, no. 4, pp. 49–56, 2000.

[4] S. Klein and J. D. anad WA Beckman, "Trnsys: A transient simulation program," Tech. Rep., 1976.

[5] B. Birdsall, W. Buhl, K. Ellington, A. Erdem, and F. Winkelmann, "Overview of the doe-2 building energy analysis program, version 2. 1d," Tech. Rep., Feb 1990.

[6] M. Wetter, "Multizone building model for thermal building simulation in modelica," in 5^{th} *International Modelica Conference*, vol. 2, 2006, pp. 517–526.

[7] Z. Lin, T. Chow, J. Liu, and K. Fong, "Cfd simulation of transient cooling in a typical hong kong office," 2000.

[8] M. Gouda and S. D. C. Underwood, "Low-order model for the simulation of a building and its heating system," *Building Services Energy Research Technology*, vol. 21, pp. 199–208, Aug 2000.

[9] S. Wang and X. Jin, "Model-based optimal control of VAV air-conditioning system using genetic algorithm," *Building and Environment*, vol. 35, no. 6, pp. 471 – 487, 2000.

[10] M. Gouda, S. Danaher, and C. Underwood, "Building thermal model reduction using nonlinear constrained optimization," *Building and Environment*, vol. 37, pp. 1255–1265, 2002.

[11] Z. Neill, S. Narayanan, and R. Brahme, "Model-based thermal load estimation in buildings," in *In SimBuild, 4th national conference of IBIPSA*, 2010.

[12] Z. Zhai and Q. Chen, "Performance of coupled building energy and cfd simulations," *Energy and buildings*, vol. 37, pp. 333–344, April 2005.

[13] H. Awbi and A. Hatton, "Natural convection from heated room surfaces," *Energy and buildings*, vol. 30, pp. 233–244, Aug 1999.

[14] S. Goyal and P. Barooah, "A method for model-reduction of non-linear thermal dynamics of multi-zone buildings," 2011, submitted for journal publication. [Online]. Available: http://plaza.ufl.edu/siddgoya/Homepage/Publications.html

[15] K. Deng, P. Barooah, P. Mehta, and S. Meyn, "Building thermal model reduction via aggregation of states," in *American Control Conference*, 2010, pp. 5118–5123.

[16] 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.

[17] M. Drton and M. Perlman, "Model selection for gaussian concentration graphs," *Biometrika 91*, pp. 591–602, 2004.

[18] "University of florida department of physics weather station." [Online]. Available: http://www.phys.ufl.edu/weather/