

EFFICIENT AND ROBUST TRAINING METHODOLOGY FOR INVERSE BUILDING MODELING

Jie Cai¹ and James E. Braun²

¹Ray W. Herrick Laboratory, Purdue University, US

²Ray W. Herrick Laboratory, Purdue University, US

ABSTRACT

This paper expands on a previous approach for inverse building modeling that utilizes a simplified state-space approach. The goal of the current effort is to provide an efficient and robust parameter training methodology, to which several elements are added. Some seasonal effects, such as variation of window transmittance at different times of the year, are taken into consideration and captured during the training process. In addition, a mixed-mode training approach is developed that allows the use of a combination of data obtained when cooling or heating is occurring with the zone temperature under control at setpoint and when the zone temperature is floating during periods of no load. Different search algorithms were tested for learning a “nearly” global optimal model. A multi-start search method was found to be robust and provide good computational efficiency and accurate results. At the end of this paper, this training methodology is implemented for a single zone case study and some results are provided.

INTRODUCTION

Forward building models, such as those employed in EnergyPlus (2011) and TRNSYS (2010), require many geometrical and physical parameters and are appropriate for system design but are not ideal for online applications, such as real-time control or fault identification. Inverse building modeling is a grey-box data driven modeling process and it typically utilizes simplified modeling approaches with lumped parameters where the parameters are learned using data collected from a specific building site. When properly trained, these types of models can provide more accurate and computational efficient load predictions for a specific building and are more appropriate for the operational phase of the building cycle. In addition to real-time control and fault identification, these types of models can be used in analyzing retrofit opportunities.

MODEL STRUCTURE AND UNKNOWN PARAMETERS

Model Structure

Previous studies (Chaturvedi, 2000) have shown that a two-node representation of a single wall provides a reasonably accurate approximation of the actual wall. Based on this starting point, a simplified structure of a

zone can be proposed using two nodes to represent each wall. A simplified whole building model that was developed by Chaturvedi and Braun (2002) is shown below in Figure 1.

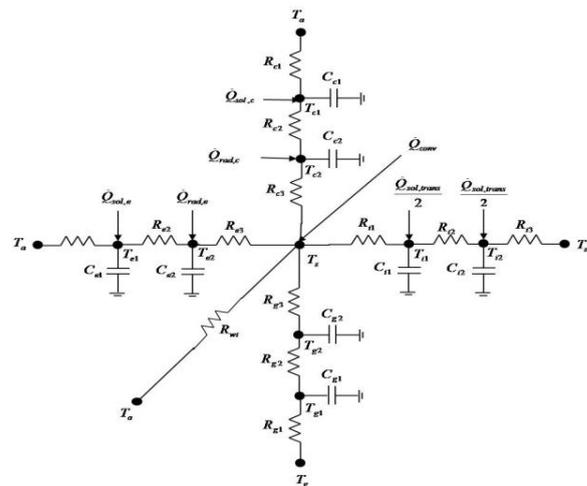


Figure 1: Thermal network for a single zone building model.

In this representation, all of exterior walls are combined into a single exterior wall with an external boundary condition that includes the total incident radiation on all wall surfaces. Solar radiation that is transmitted through windows is assumed to be absorbed equally on two sides of an interior wall presentation. The interior wall also captures the effects of floors between stories of a multi-story building. An additional ground element is included to capture ground coupling dynamics. Internal radiative gains are assumed to be distributed with an even flux to walls and ceiling, whereas convective internal gains go directly to the zone air. A pure resistance is included to capture the effects of heat transfer across low-mass elements, such as windows or due to infiltration.

Applying an energy balance to each node in the network, a state-space representation can be established for this simplified model structure:

$$\frac{d\mathbf{x}_b}{dt} = \mathbf{A}_b\mathbf{x}_b + \mathbf{B}_b\mathbf{u}_b \quad (1)$$

$$\mathbf{Y}_b = \mathbf{C}_b\mathbf{x}_b + \mathbf{D}_b\mathbf{u}_b \quad (2)$$

where \mathbf{Y}_b is the output, which can be either cooling load (zone air temperature is input) or zone air temperature (cooling load is input).

For the case of cooling load being the output, the state and input vectors are:

$$\mathbf{x}_b^T = [T_{c1} T_{c2} T_{e1} T_{e2} T_{g1} T_{g2} T_{i1} T_{i2}]$$

$$\mathbf{u}_b^T = [T_z T_a T_g \dot{Q}_{sol,c} \dot{Q}_{sol,e} \dot{Q}_{rad,c} \dot{Q}_{rad,e} \dot{Q}_{sol,trans} \dot{Q}_{conv}]$$

For the case of zone air temperature being the output, the state and input vectors are:

$$\mathbf{x}_b^T = [T_{c1} T_{c2} T_{e1} T_{e2} T_{g1} T_{g2} T_{i1} T_{i2} T_z]$$

$$\mathbf{u}_b^T = [\dot{Q}_b T_a T_g \dot{Q}_{sol,c} \dot{Q}_{sol,e} \dot{Q}_{rad,c} \dot{Q}_{rad,e} \dot{Q}_{sol,trans} \dot{Q}_{conv}]$$

Unknown Parameters

The set of parameters to be estimated can be denoted as:

$$\theta = \{ \theta_1, \theta_2 \}$$

where

$$\theta_1 = [C_{e1} C_{i1} C_{c1} C_{g1} R_{e1} R_{e2} R_{i1} R_{i2} R_{c1} R_{c2}]$$

$$R_{g1} R_{g1} R_{g2} C_{e2} C_{i2} C_{c2} C_{g2} R_{e3} R_{i3} R_{c3} R_{c3}]$$

$$\theta_2 = [Q_{int,gain} R_{conv} a_{trans} b_{trans}]$$

The first group θ_1 consists of all resistances of both the walls and the window, and capacitance values of the walls in the thermal network in Figure 1. The resistances and capacitances determine the entries of matrices \mathbf{A}_b , \mathbf{B}_b , \mathbf{C}_b , and \mathbf{D}_b of Equations (1) and (2). The second group θ_2 contains parameters that are used for all other purposes. The first two parameters in the 2nd group are used for the rate of internal gains ($Q_{int,gain}$) during the occupied period and the ratio of the convective component (R_{conv}) to the total internal gains. When the model is trained using data generated from simulation tools like TRNSYS, the internal gains can be exported as input for the model and these two parameters are just place keepers. But when using actual field data for training, it is not possible to measure the actual internal gains. In this case, the internal gains and their effect should be captured in the training and the quantities are represented by these two parameters. In our case, an assumption was made for this study that during the unoccupied period, the internal gains are zero while in the occupied period they are a constant gain of $Q_{int,gain}$. This internal gain

splits between convective and radiative components at a fixed ratio, which is denoted by $R_{conv} / (1 - R_{conv})$.

The other two parameters (a_{trans} , b_{trans}) are used in modeling the variation of transmittance versus the incident solar angle for the window. The details are illustrated in the following section.

Transmittance variation

Window transmittance can be classified into beam and diffuse radiation transmittance (Arasteh et al., 2009). Beam transmittance varies with incidence angle while diffuse transmittance is constant. Based on typical trends for window transmittance with incidence angle, the following correlations were developed:

$$T_{beam} = T * F = T * (1 - a_{trans} \alpha^n)$$

$$T_{diff} = T * b_{trans}$$

where T_{beam} and T_{diff} represent beam and diffusive transmittance, α is solar incidence angle and n is correlation order. a_{trans} and b_{trans} are correlation coefficients, which are also parameters to be estimated. T is the total transmittance, which is taken as the solar heat gain coefficient value (SHGC) associated with each window model for simplicity. The estimation of a_{trans} and b_{trans} is embedded in the whole training process and some results of tested cases are shown in Figure 2.

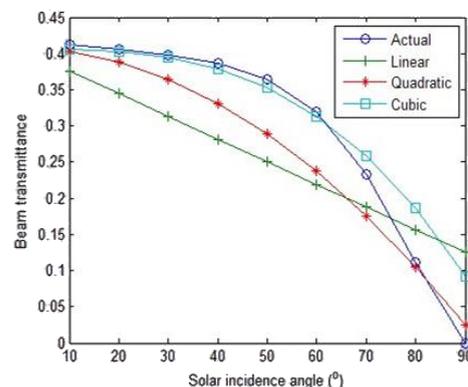


Figure 2: Beam transmittance versus solar incidence angle.

The window model used as the baseline for Figure 2 was INS2_KR_3 from the TRNSYS window library and the estimation is embedded in the single zone model in the case study section. The blue curve in Figure 2 was generated from the TRNSYS window model, whereas the green, red and light blue curves correspond to linear, quadratic and cubic correlations, respectively, with the following parameter values:

$$a_{trans} = 0.0077, n = 1$$

$$a_{trans} = 0.000116, n = 2$$

$$a_{trans} = 0.00000106, n = 3$$

The estimated value of b_{trans} was within the range of 6.2 to 6.5 for all three cases. The linear correlation was found to be accurate enough and higher order correlations did not improve overall system model results significantly.

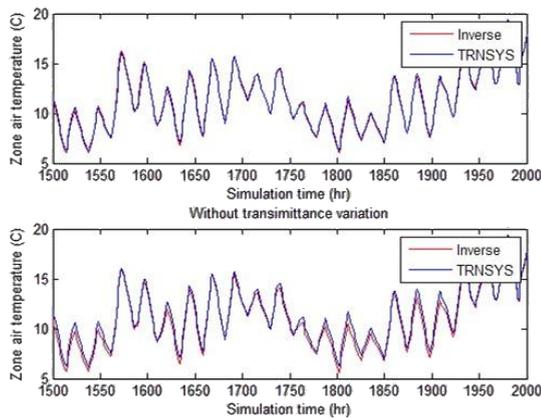


Figure 3: Model performance comparison with and without transmittance variation.

Figure 3 shows comparisons of predicted floating temperature where summer data were used for training (hour 5300 to 5500 within the year) and winter data were used for testing (hour 1500 to 2200). A description of the single zone case study building is given in a later section. The RMS errors during the test period were 0.501°C for the case where a fixed transmittance was assumed and 0.185°C when transmittance varied using linear correlation with respect to solar angle. The performance was improved significantly.

PARAMETER ESTIMATION

Problem Formulation

For a fixed set of parameter values (θ is fixed), the methodology of Seem et al. (1989) can be used to solve the state-space representation, and to predict cooling load or zone air temperature of the building as a transfer function of the input and state variables. The time step of the transfer function is assumed be one hour. The solution for the output is given as:

$$y_{b,k} = \left(\sum_{j=0}^8 \mathbf{S}_{b,j} \mathbf{u}_{b,k-j} \right) - \left(\sum_{j=1}^8 u e_{b,j} y_{b,k-j} \right)$$

where

$y_{b,k}$ = output (either $\dot{Q}_{b,k}$ or $T_{b,k}$) at time k ;

$\mathbf{S}_{b,j}$ = series of row vectors containing transfer function coefficients for the input matrix in the building state-space representation;

$e_{b,j}$ = scalar transfer function coefficients for past histories of output.

The calculation of the coefficients $\mathbf{S}_{b,j}$ and $e_{b,j}$ from matrices \mathbf{A}_b , \mathbf{B}_b , \mathbf{C}_b and \mathbf{D}_b is outlined by Seem et al. (1989).

For a specific set of parameter values, the performance of the model can be evaluated in terms of how well the output (or prediction) matches the actual data (baseline). A commonly used criterion for the deviation of prediction from baseline is least square error:

$$J_{b,\theta} = \sqrt{\frac{\sum_{k=1}^{N_{train}} (y_{b,k} - y_{actual,k})^2}{N_{train} - 1}}$$

where

$J_{b,\theta}$ = optimization regression cost function at point θ ;

$y_{actual,k}$ = baseline output at time step k ;

N_{train} = training duration in hours.

So the formulation of our parameter estimation problem can be written as:

$$\theta^* = \arg \min_{\theta \in \Omega} (J_{b,\theta})$$

where Ω is the search region for the parameter values in the estimation process. This search region is determined according to the information available about the zone. And the information can be obtained via several means, such as by requesting a survey from the building administrator or by looking at the blueprint of the building. Generally speaking, the less information we have about the building, the bigger the search region is for the optimization. As search region grows, it becomes more likely that the estimation process will converge to a local optimum. So some preprocessing is necessary in obtaining a good initial guess to improve the chances of obtaining a global optimal point.

Global-Local Search

In the previous work by Chaturvedi and Braun (2002), a global-local search scheme was adopted (see Figure 4) for parameter estimation. A global phase was applied in order to obtain a global optimal estimation of the parameter values, and a systematic search algorithm developed by Aird and Rice (1977) was implemented. It is similar to grid search but more generalized in the sense that any number of points can be generated in the

search space with the property that the dispersion of the generated set is minimized. At each of these generated points, the least square error is evaluated and the point with a smallest least square error is chosen as an initial guess for the local search.

In the local search phase, the Levenburg-Marquardt algorithm (Madsen et al., 2004) is used. This LM algorithm interpolates between the Gauss-Newton and the gradient descent method but it shows more robustness than the Gauss-Newton algorithm.

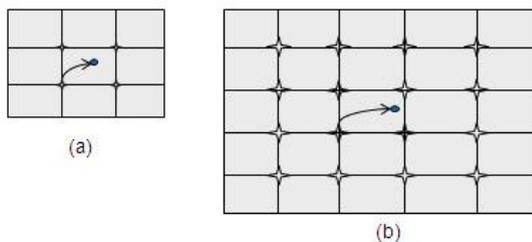


Figure 4: (a) Small search region. (b) Large search region, in which the number of point evaluations is increased to maintain the same level of gridding.

Multi-Start Search

When very limited information about a building is available, the search region becomes large and it is not feasible to use the global-local search scheme. In this situation, a multi-start search scheme would be more suitable since this type of methods is often applied for large-scale estimation problems (Aster et al., 2005).

In a multi-start search process, several points are generated pseudo-randomly as initial guesses for regression and the regression is performed for each point. The solution with the smallest least square error is chosen to be the final parameter values.

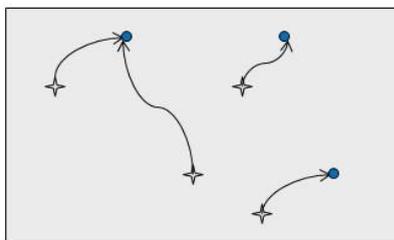


Figure 5: Multi-start search scheme.

Some comparisons of results for the two search schemes were carried out for the single zone case study building described in the next section with results given in Table 1

Table 1: Comparison of performances of global-local and multi-start search algorithms

	Global-local Search	Multi-start Search
Search region lower bound:	10% of nominal parameter values	5% of nominal parameter values
Search region upper bound:	300% of nominal parameter values	400% of nominal parameter values
Training time:	>30 min	45s
Testing case RMSE:	>10%	0.24%

Numbers in the first two rows of Table 1 are lower and upper bounds on the search region for parameters expressed as a percentage of nominal parameter values that were estimated from zone descriptions. The large search region was chosen to represent a case where very limited information is available on the building. When detailed zone information is available, more reasonable upper and lower bounds might be 130% and 70% of nominal values. Even though the chosen search region was larger for the multi-start search method, this method provided more accurate zone air temperature predictions with significantly less computational time for training. The multi-start search method is more robust and efficient when the parameter search region is large.

Mixed Training Mode

There are two modes in which models can be trained. In one mode, the training process takes zone air temperature as input and predicts cooling load, while in the other mode, the input and output are switched. In a real operational phase, the zone air temperature may be floating sometimes when the cooling equipment is off, (e.g., during the unoccupied period) and at other times it may be under control (e.g., positive cooling load in the occupied period). The fraction of the time that there is a cooling load changes with seasons and control strategies. When the zone temperature is under control, there are no dynamics in the zone temperature output so it is better to train the model using cooling load as the output. Conversely, when the cooling is off and the temperature is floating, temperature is the preferred output for training. Based on this analysis, a mixed training mode is proposed whereby the training mode switches from cooling or zone temperature as an output according to which output has better dynamics. Figure 6 shows the mechanism of this training process.

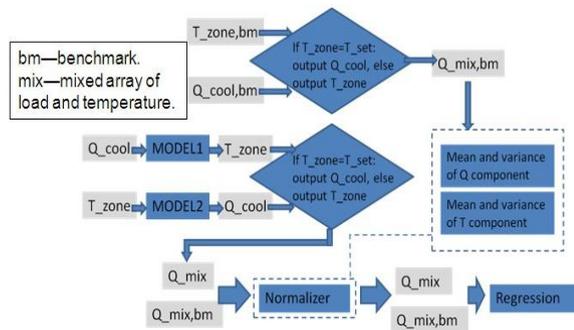


Figure 6: Mechanism of mixed training process.

Table 2 provides some comparisons between training approaches using global-local search scheme for a test period same as the one in the case study section. In this table, the variable P represents the percentage of the training period where the zone temperature was under controlled conditions. For the first case, P had a small value and the zone air temperature was floating most of the time. In this case, using temperature as the output for training provided the best performance. In the second case with an even mix of floating and controlled conditions, this mixed training was most advantageous since the temperature and cooling load dynamics were both well captured. For the third case where the zone air temperature is under control most of the time, training using temperature and training using load both have good performance.

Table 2: Performance of different training mode approaches in terms of prediction errors (P is percentage of the training period where the zone temperature was under controlled conditions)

P=23%	Mixed training	Training using load	Training using temperature
Zone Temperature (RMSE °C)	0.33	1.05	0.298
Zone Load (% RMSE)	2.73	3.0	2.70

P=65%	Mixed training	Training using load	Training using temperature
Zone Temperature (RMSE °C)	0.207	0.38	0.273
Zone Load (% RMSE)	2.82	4.03	3.42

P=86%	Mixed training	Training using load	Training using temperature
Zone Temperature (RMSE °C)	0.913	0.313	0.283
Zone Load (% RMSE)	6.65	2.51	2.56

SINGLE ZONE CASE STUDY

Case Study Definition

The inverse modeling and training algorithms described in the preceding sections were applied to a single zone case study where the baseline training data was generated using TRNSYS. Some details of the case study parameters are listed below (see Figure 7).

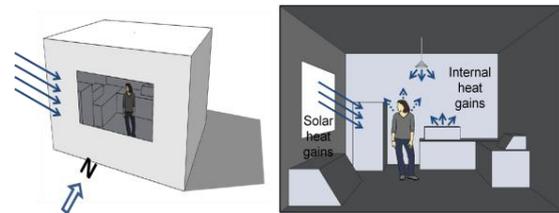


Figure 7: Single zone case study.

Geometry of zone:

- Zone size: 10 by 10 by 10 (m).
- Window size: 7 by 7 (m), on south wall.
- Internal wall size: 7 by 7 (m).

Construction information:

- Wall (including all walls, ceiling and ground) material: concrete.
- Wall (including all walls, ceiling and ground) thickness: 0.2 (m).
- Window: INS2_KR_3 from library WINDOW 4.1; insulating glazing with Krypton as gas fill.

Other information:

- Weather location: Madison, Wisconsin.
- Control strategy: precool.
- Training period: hour 5300 to 5500 from TMY2.

Case Study Results

Testing results for the period from 5080th to 5250th hours of TMY2 data are shown in Figure 8. One plot shows prediction of zone air temperature for the test period when cooling load was provided as an input and the other curve gives predictions of cooling load given

zone air temperature. The colors black and red indicate predicted data from the inverse model and output from TRNSYS, respectively. It can be seen that the prediction matches the TRNSYS outputs very well visually. The overall RMS error is 0.37°C for the zone air temperature prediction and 2.8% for cooling load prediction, both in the least square sense.

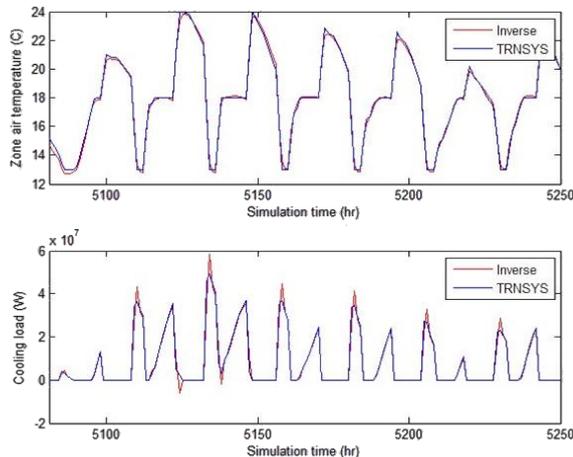


Figure 8: Case study results.

FUTURE WORK

At the current stage, the inverse modeling training approach has only been applied to simulated data. The next step will be to apply it to an existing building that includes multiple zones. When the number of zones is large, a decentralized parameter estimation technique may be required because of computational issues associated with determining parameters in a large dimensional space.

REFERENCES

- Arasteh, D., Kohler, C. and Griffith B., 2009, Modeling Windows in Energy Plus with Simple Performance Indices, Lawrence Berkeley National Laboratory.
- Aird, T.J. and Rice, J.R., 1977, Systematic Search in High Dimensional Sets, SIAM Journal on Numerical Analysis, Vol.14, pp. 296-312.
- Braun, J.E. and Chaturvedi, N., 2002, An Inverse Gray-box Model for Transient Building Load Prediction, HVAC&R Research, Vol. 8, No. 1, pp. 73-99.
- Chaturvedi, N., 2000, Analytical Tools for Dynamic Building Control, Masters Thesis, Herrick Laboratories, West Lafayette, Indiana.

K. Madsen, H.B. Nielsen and O. Tingleff, 2004, Methods for Nonlinear Least Squares Problems, Informatics and Mathematical Modeling, Technical University of Denmark.

Seem, J.E., S.A. Klein, W.A. Beckman, and J.W. Mitchell, 1989, Transfer Functions for Efficient Calculations of Multi Dimensional Heat Transfer, Journal of Heat Transfer-Transactions of the ASME 111(1):5-12.

Aster, R.C., Borchers, B. and Thurber, C.H., 2005, Parameter Estimation and Inverse Problems.

TRNSYS 17, 2010, Solar Energy Laboratory, University of Wisconsin-Madison.

EnergyPlus 7.0, 2011, <http://apps1.eere.energy.gov/buildings/energyplus/>, US Department of Energy.

ACKNOWLEDGEMENT

This work was supported by the Department of Energy through the Energy Efficient Buildings Hub.

NOMENCLATURE

$\dot{Q}_{sol,e}$ = rate of solar radiation incident on the lumped external walls

$\dot{Q}_{sol,c}$ = rate of solar radiation incident on the roof (ceiling)

$\dot{Q}_{sol,trans}$ = rate of solar radiation transmitted through windows to the inside

$\dot{Q}_{rad,c}$ = radiative internal gains per unit time transmitted to the ceiling

$\dot{Q}_{rad,e}$ = radiative internal gains per unit time transmitted to the external walls

\dot{Q}_{conv} = convective internal gains per unit time

T = temperature

R = resistor

C = capacitor

e = external wall

i = internal wall

c = roof (ceiling)

g = ground

z = zone air

a = ambient air

wi = windows

mix = mixed array of load and temperature in mixed training mode

bm = benchmark data (training data)