Estimation of Building Parameters Using Simplified Energy Balance Model and Metered Whole Building Energy Use

Hiroko Masuda[1] and David E. Claridge[2]

[1, 2] *Energy Systems Laboratory, Texas Engineering Experiment Station*
Texas A&M University, College Station, TX. U.S.A.

Email: hirokomasuda@tees.tamus.edu

Abstract:
This paper presents and evaluates an indirect data-driven method to estimate influential building parameters: air exchange rates and overall heat transfer coefficients of building envelopes from the separately metered energy use for electricity, cooling and heating and weather data using multiple linear regression models based on the simplified steady-state energy balance for a whole building. Two approaches using different response variables: the energy balance load (E_{BL}) and the building thermal load (Q_B) and the use of monthly and daily interval data are evaluated using the synthetic data and the measured data from three large dormitory buildings. Although this method is not expected to replace actual measurement, easy and fast access to the influential building parameters allows new applications such as in preliminary investigation for energy conservation projects.

Keywords:
energy data analysis, parameter identification, data-driven models, energy balance

1. Introduction

Air exchange rates and overall heat loss coefficient of the building envelope have significant influences on the heating and cooling loads, but direct measurement of these parameters for a whole building scale in operating buildings requires considerable time and labor. This paper presents an indirect data-driven method to estimate the building parameters for operating buildings without direct measurement.

In data-driven approaches, the parameters are statistically estimated based on the relationship between input and output data. To obtain parameter estimates that allow direct physical interpretation, one needs to formulate models based on physical principles such as energy balance. Transient or steady-state models can be chosen depending on the required resolution of the model prediction, available data period and intervals, etc. This paper focus on the steady-state models using daily or monthly interval data to minimize transient effects due to building thermal mass and to average out the variations of ventilation rate and internal loads in a day so that the parameters can be reasonably constant over the modeling period.

Earlier studies on steady-state building parameter estimations generally use heating or cooling energy consumption data as a response variable in the regression models. However, using these models, simultaneous cooling and heating commercial buildings can cause misleading parameter estimations (Rabl and Rialhe, 1992). For commercial buildings, Reddy et al. (1994) explored the method to infer basic building parameters using the variable called the building thermal load Q_B. The Q_B variable is calculated from the whole building cooling and heating energy use, and the mixed hot and cold air streams in the HVAC systems in the building will be canceled out. Deng (1997) has developed a data-driven method to estimate overall heat loss coefficient and ventilation parameter using the Q_B model, and tested it with synthetic data. The parameter estimates are found to be accurate when daily data over an entire year are used, and biases due to multicolinearity—correlations between explanatory variables can be minimized using the estimation technique they developed, so called multi-step linear regressions. Another variable called the energy balance load E_{BL} has been proposed by Shao and Claridge (2006) as a part of energy data screening. The E_{BL} variable is similar to Q_B, but it includes the heat load from the electrical energy use. The E_{BL} variables plotted as a function of the outside air temperature shows a linear pattern that is unique to each building and faulty data can be detected as outliers visually or statistically using empirical models (Baltazar et al. 2007 and 2012).

In this paper, the method to estimate building parameters using the E_{BL} variable will be presented along with the method using the Q_B variable, and the estimation results are compared. Synthetic data from a computer simulation are used to evaluate the estimation, and to assess the application of the method to actual buildings, comparisons of measured outside air flow rates and the estimated values are presented. The use of newly introduced reference parameter T_{in}^* is discussed to alert unreliable parameter estimations.

Expected uses of the estimated parameters include: preliminary information for initial investigations in building energy optimization projects, reference input values in building energy simulations for existing buildings as a supplement to calibration procedures such as Claridge et al. (2003), and detection of operational changes with continuous monitoring. The automated application of the method as a part of energy information systems may be suitable for the facilities that require monitoring of large amount of buildings at the same time such as college campuses.

2. Formulation of models

2.1. Definition of E_{BL} and Q_B

Unified mathematical expressions for the parameters of the models using E_{BL} and Q_B variables are presented first. A system including the entire building is chosen as a control volume, and the boundary is set right outside the building exterior surfaces. The net change in the total energy of the control volume ΔE_{CV} is equal to the difference between the total energy entering and leaving the system. That is,

$$\begin{aligned}\Delta E_{CV} &= E_{entering} - E_{leaving} \\ &= Q_{air} + Q_{cond} + Q_{sol} + Q_{occ} + Q_E - E_C + E_H \end{aligned} \qquad (1)$$

where Q_{air}, Q_{cond}, Q_{sol}, Q_{occ}, and Q_E are building heat load components from air exchange, conduction through exterior surfaces, solar insolation, occupants, and electricity energy consumed in the building, respectively. E_C, and E_H are separately metered whole building energy use of cooling and heating. When the time scale under study is long enough to diminish the thermal lag effect and the indoor air thermal condition is maintained constant, the system can be considered as a quasi-steady state, and the left hand side of Eq. (1) yields zero. The energy balance load E_{BL} is defined as (Shao, 2006):

$$\begin{aligned} E_{BL} &= Q_E - E_C + E_H \\ &= fE_E - E_C + E_H \\ &= -Q_{air} - Q_{cond} - Q_{sol} - Q_{occ} \end{aligned} \tag{2}$$

where E_E is the metered whole-building non-cooling electricity use. The multiplicative factor f represents a fraction of E_E which turns into the heat load ($0 \leq f \leq 1$). The factor f is not measurable but presumed to be fairly high. In practice, the available whole building level of electricity consumption is used for Q_E to calculate E_{BL}. In this case, Eq. (2) is re-written as:

$$\begin{aligned} E_{BL} &= E_E - E_C + E_H \\ &= -Q_{air} - Q_{cond} - Q_{sol} - Q_{occ} + (1-f)E_E. \end{aligned} \tag{3}$$

The electricity energy use which does not turn into the space heat load may increase the E_{BL}. Meanwhile, the building thermal load Q_B is defined as (Reddy et. al. 1994):

$$\begin{aligned} Q_B &= E_C - E_H \\ &= Q_{air} + Q_{cond} + Q_{sol} + Q_{occ} + Q_E \\ &= Q_{air} + Q_{cond} + Q_{sol} + Q_{occ} + fE_E. \end{aligned} \tag{4}$$

Deng (1997) and Reddy et al. (1999) introduced two multiplicative correction factors: k_s and k_l. The k_s is a fraction of internal sensible loads to measured electricity use of lights and equipment E_{LE} and the k_l is a fraction of internal latent load to the total internal sensible load which appears only when latent load exists. If all the internal loads are from the occupants and lights and equipment, this relationship is written as

$$Q_{occ} + fE_E = E_{LE} k_s (1 + k_l X) \tag{5}$$

where the indicator variable X is 1 when the latent load exists ($W_{oa} > W_{in}$) and 0 otherwise. Then the expression of Q_B becomes

$$Q_B = Q_{air} + Q_{cond} + Q_{sol} + E_{LE} k_s (1 + k_l X). \tag{6}$$

2.2. Key parameters

The problem will be simplified as the same manner as in Reddy et al. (1994 and 1999), Deng (1997), and Shao (2006). The assumptions for the simplified E_{BL} and Q_B models using daily or monthly resolution data can be summarized as follows.

1. Indoor air temperature T_{in} is constant.
2. Indoor humidity W_{in} does not exceed 0.01 kg/kg. No humidification is applied.
3. Overall heat loss coefficient and air exchange rate are constant.
4. No economizer or heat recovery device is used.
5. Building total solar load can be expressed as a linear function of the outside air temperature T_{oa}.
6. Occupancy load is overall constant.
7. Transient effect is negligible.

Based on these assumptions, Q_{air}, Q_{cond}, and Q_{sol} are expressed in the simplified steady-state load models as presented in Eq. (7), (8), and (9).

$$Q_{air} = m_v c_p (T_{oa} - T_{in}) + m_v h_v X (W_{oa} - W_{in}) \tag{7}$$

$$Q_{cond} = UA_s (T_{oa} - T_{in}) \tag{8}$$

$$\begin{aligned} Q_{sol} &= a_{sol} + b_{sol} T_{oa} \\ &= a'_{sol} + b_{sol}(T_{oa} - T_{in}) \end{aligned} \tag{9}$$

where a and b are constants. W_{in} = 0.01 kg/kg and the indicator variable X is 1 when (W_{oa}>W_{in}) and 0 otherwise. By inserting these into Equations (3) and (6), the multiple linear regression models for E_{BL} and Q_B are derived as

$$E_{BL} = \beta_0 + \beta_T T_{oa} + \beta_W X (W_{oa} - W_{in}) + \varepsilon \tag{10}$$

$$Q_B = \beta_0 + \beta_{sens} E_{LE} + \beta_{lat} X E_{LE} + \beta_T T_{oa} + \beta_W X (W_{oa} - W_{in}) + \varepsilon \tag{11}$$

where ε is a random error. The mathematical expressions for each regression parameter are presented in Table 1.

Table 1: Mathematical expressions for regression model parameters

Regression parameter	$\boldsymbol{E_{BL}}$	$\boldsymbol{Q_B}$
β_0	$(UA_s + m_v c_p + b_{sol})T_{in} - Q_{occ}$ $-a'_{sol} + (1-f)E_E$	$-(UA_s + m_v c_p + b_{sol})T_{in} + a'_{sol}$
β_T	$-(UA_s + m_v c_p + b_{sol})$	$UA_s + m_v c_p + b_{sol}$
β_W	$-m_v h_v$	$m_v h_v$
β_{sens}	Not available	k_s
β_{lat}	Not available	$k_s k_l$

From these expressions, the building parameters and the uncertainties are deduced as in Table 2. The overall heat loss coefficient estimated from the regression models cannot separate out the solar effect. To differentiate it from the U for the temperature difference between indoor and outdoor air, we use U^* which is defined as $U^*A_s=UA_s+b_{sol}$. The variable T^*_{in} is introduced as a reference parameter which is associated with the indoor air temperature T_{in} and resembles the balance point temperature (ASHRAE 2009). The physical interpretation of this parameter changes depending on the explanatory variables included in the regression model, which will be discussed later along with the estimation results.

3. Data and procedures

3.1. Synthetic data

The commercial reference building model for existing large office buildings constructed in or after 1980 in the climate zone with the representative city of Houston, TX (DOE 2010) is used to generate synthetic data using EnergyPlus simulation software. The building has 12 stories above ground and a basement, and the total conditioned area of 46,320.38 m^2 (498,588 ft^2). Each above-grade floor has 5 zones: north, east, south, and west perimeters, core, and plenum. Each floor has a single duct VAV system with reheat terminals, and the building does not use economizer.

Table 2: Equations to calculate building parameters and the uncertainties from the regression estimates and standard errors

Building parameter	E_{BL}	Q_B	Uncertainty
m_v	$-\hat{\beta}_W / h_v$	$\hat{\beta}_W / h_v$	$\Delta\hat{\beta}_W / h_v$
$U^* A_s$	$-\hat{\beta}_T + \hat{\beta}_W c_p / h_v$	$\hat{\beta}_T - \hat{\beta}_W c_p / h_v$	$\sqrt{(\Delta\hat{\beta}_T)^2 + (\Delta(\hat{\beta}_W \cdot c_p / h_v))^2}$
T_{in}^*	$-\hat{\beta}_0 / \hat{\beta}_T$	$-\hat{\beta}_0 / \hat{\beta}_T$	$\left(\frac{\hat{\beta}_0}{\hat{\beta}_T}\right)\sqrt{\left(\frac{\Delta\hat{\beta}_0}{\hat{\beta}_0}\right)^2 + \left(\frac{\Delta\hat{\beta}_T}{\hat{\beta}_T}\right)^2}$
k_s	Not available	$\hat{\beta}_{sens}$	$\Delta\hat{\beta}_{sens}$
k_l	Not available	$\hat{\beta}_{lat} / \hat{\beta}_{sens}$	$\left(\frac{\hat{\beta}_{lat}}{\hat{\beta}_{sens}}\right)\sqrt{\left(\frac{\Delta\hat{\beta}_{lat}}{\hat{\beta}_{lat}}\right)^2 + \left(\frac{\Delta\hat{\beta}_{sens}}{\hat{\beta}_{sens}}\right)^2}$

Note: the delta means the standard error

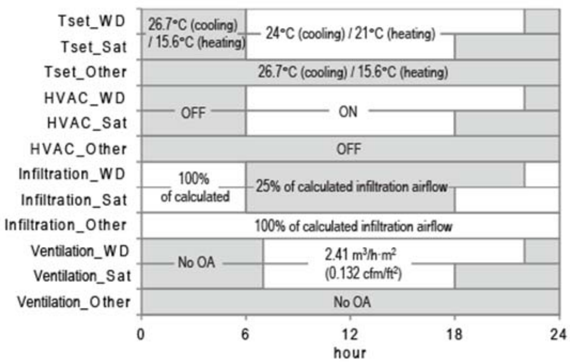

Fig. 1. System schedules for the as-is case

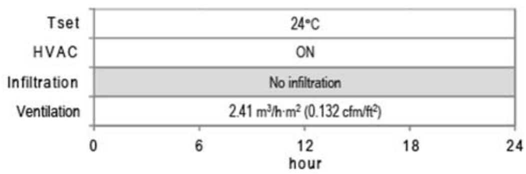

Fig. 2. System schedules for the ideal cases

Building operation in the original input file has three schedule patterns for weekday (WD), Saturday (Sat), and Sunday and holidays (Other) as shown in Fig. 1. During unoccupied hours, the HVAC systems are turned off until the zone temperatures exceed the set point temperatures; different set points are defined for cooling or heating and occupied or unoccupied hours. Another input file was prepared by modifying the schedules as in Fig. 2 to generate ideal data for parameter identifications. Then, the three sets of synthetic data were generated as in Table 3. Figure 3 shows the daily energy uses for electricity (lights, equipment, and fans), cooling, and heating, and Fig. 4 shows the E_{BL} and Q_B variables evaluated from these energy uses plotted versus the daily average temperature for the as-is case. Note that the signs of the Q_B plots are switched for the ease of visual comparison. Figures 5 and 6 show the same plots for the ideal w/o solar case.

Table 3: Measured values of outside air flow rates and energy data periods of E_{BL} and Q_B for three dormitory buildings

Case designation	**System schedules**	**Weather data**
As is	Fig. 1	TMY2 for Houston, TX
Ideal w/ solar	Fig. 2	TMY2 for Houston, TX
Ideal w/o solar	Fig. 2	Modified Houston TMY2 (solar insolation = 0)

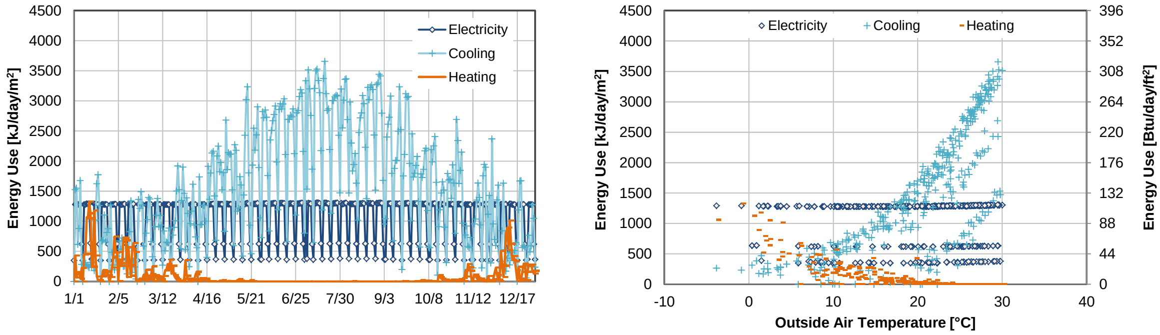

Fig. 3. Whole building daily energy uses for electricity, cooling, and heating per unit conditioned floor area for the as-is case. Time series plot is in the left and scatter plot versus daily average outside air temperature is in the right.

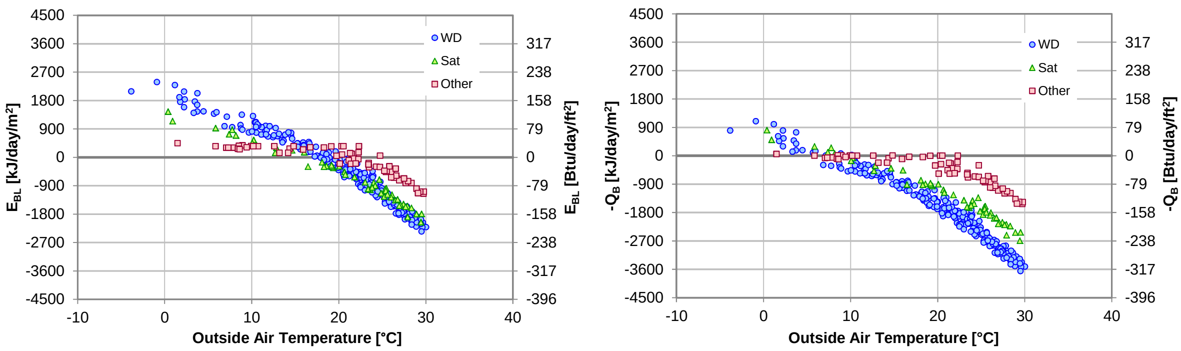

Fig. 4. E_{BL} and Q_B per unit conditioned floor area in the as-is case plotted versus daily average outside air temperature

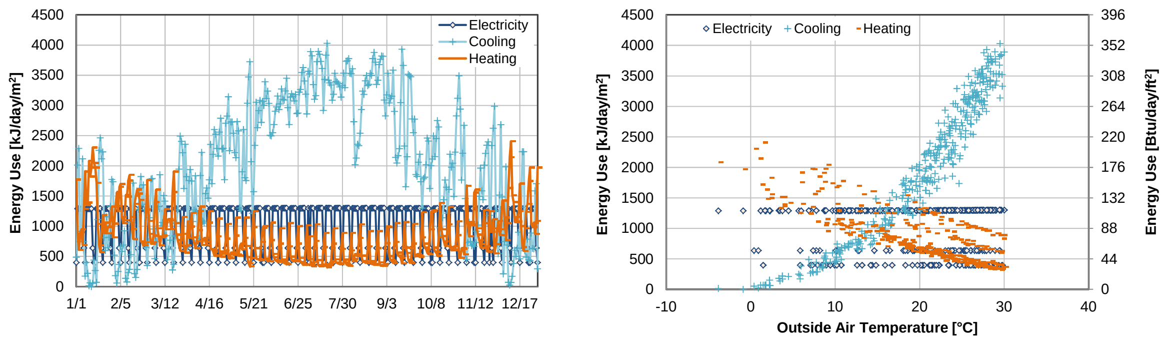

Fig. 5. Whole building daily energy uses for electricity, cooling, and heating per unit conditioned floor area for the ideal w/o solar case. Time series plot is in the left and scatter plot versus daily average outside air temperature is in the right.

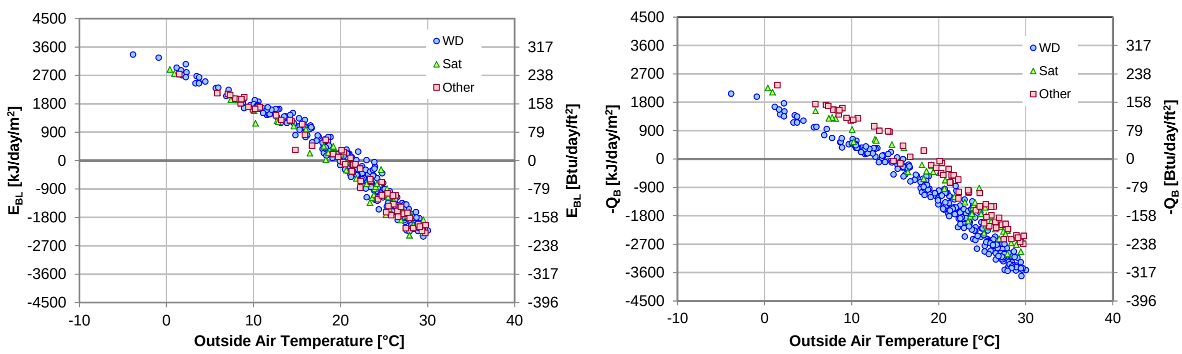

Fig. 6. E_{BL} and Q_B per unit conditioned floor area in the ideal w/o solar case plotted versus daily average outside air temperature.

3.2. Data from actual buildings

The whole building electricity, chilled water, and heating hot water energy use data are collected from the three dormitory buildings which have dedicated outdoor air systems (DOAS). The HVAC systems are operated continuously in these buildings. The outside air temperature and humidity ratio are obtained and calculated from the quality controlled local climatological data (QCLCD) for College Station, TX (NCDC 2012). The outside air flow rate has been measured at the OAHUs on 4/24/2012. The measured values of the building total outside air flow rate and the basic information on the available E_{BL} and Q_B data for these buildings are given in Table 4.

Table 4: Measured values of outside air flow rates and energy data periods of E_{BL} and Q_B for three dormitory buildings

Bldg Symbol	Gross floor area	Measured outside air flow rate on 4/24/2012	E_{BL} and Q_B data		
			Available energy use data period	No. of daily observations	No. of monthly observations
HS	69668 ft^2 (6472.4 m^2)	8779 cfm (14916 m^3/h)	7/1/2011–6/30/2012	320	12
MF	62156 ft^2 (5774.5 m^2)	10025 cfm (17033 m^3/h)	9/1/2011–6/30/2012	267	10
HB	62156 ft^2 (5774.5 m^2)	7750 cfm (13167 m^3/h)	7/21/2011–6/30/2012	329	12

3.3. Estimation procedure

Total of 14 sets of data listed in Table 5 were prepared, and the E_{BL} and Q_B variables were calculated for the each set. The data for the as-is case were grouped into three day types, because the parameters vary between those. The models have been estimated with the statistics software R (R Core Team, 2011). The electricity use variable E_{LE} is not included in the daily interval Q_B models for the separated as-is cases, because the daily electricity use for lights and equipment from the simulation is perfectly constant in the each day-type data, and the parameter estimates becomes zero. The variable XE_{LE} has been removed from the daily Q_B models for the three dormitories and from all the monthly Q_B models, because, when included, the direction of the effects becomes opposite from the physical response and/or the estimates are not statistically significant. The explanatory variable terms included in each final model are given in Table 5.

To detect the level of multicolinearity, the Variance Inflation Factors (VIF) have been calculated for each data set. The VIF is defined as:

$$\mathrm{VIF} = \frac{1}{1-\mathrm{R}_{\mathrm{i}}^2} \tag{12}$$

where $\mathrm{R_i}^2$ is the multiple coefficient of determination between the *i*-th explanatory variable and all of the other explanatory variables in the regression equation. The exact value of VIF at which muticolinearity is declared depends on the individual investigator. Some use a value of 5 and others 10 (Haan, 2002).

Table 5: Data sets used in the analysis and the explanatory variable terms included in the regression models. The checked terms are included.

Dataset	Explanatory variable terms included in the regression models					
	E_{BL}		Q_B			
	T_{oa}	$X(W_{oa}$-$W_{in})$	T_{oa}	$X(W_{oa}$-$W_{in})$	E_{LE}	XE_{LE}
Daily interval						
Ideal w/ solar	✓	✓	✓	✓	✓	✓
Ideal w/o solar	✓	✓	✓	✓	✓	✓
As is (WD)	✓	✓	✓	✓		
As is (Sat)	✓	✓	✓	✓		
As is (Other)	✓	✓	✓	✓		
HS (Jul–Jun)	✓	✓	✓	✓	✓	
MF	✓	✓	✓	✓	✓	
HB	✓	✓	✓	✓	✓	
Monthly interval						
Ideal w/ solar	✓	✓	✓	✓	✓	
Ideal w/o solar	✓	✓	✓	✓	✓	
As is	✓	✓	✓	✓	✓	
HS (Jul–Jun)	✓	✓	✓	✓	✓	
MF	✓	✓	✓	✓	✓	
HB	✓	✓	✓	✓	✓	

4. Results and Discussions

4.1. Evaluation using Synthetic Data

The air exchange rates m_v converted into volumetric flow rate, overall heat loss coefficients U^* estimated from the daily interval synthetic data are compared to the assumed true values in Fig. 7. The estimates for the temperature parameter β_T are also presented for reference. The signs of the β_T estimates for E_{BL} and Q_B models are opposite, and the absolute values $|\beta_T|$ are used for comparison.

Overall, E_{BL} and Q_B models have consistent parameter estimates for the daily interval synthetic data sets. In the ideal cases, despite the solar insolation effect, the m_v is estimated reasonably accurate within 10%. The solar insolation increased the $|\beta_T|$ estimates by 9.1% (Q_B) and 6.8% (E_B) and decreased the m_v estimates by 6.8% (Q_B) and 7.9% (E_B), which directly resulted in the overestimation of U^*. It should be reminded that the true value of the overall heat loss coefficient in Fig. 7 is for U which does not include solar effect and smaller than U^*, and the overestimation includes this difference.

For the WD and Sat day types in the as-is case, the parameters are estimated fairly well and comparable to the ideal cases, nevertheless these simulation models have some exceptions from the model assumptions. The $|\beta_T|$ estimates in the WD and Sat day types are seemingly as good as in the ideal cases, however, we should be cautious of this result. The T_{in} decreases with the T_{oa} in the as-is case because of the set point and system operation schedules, which may decrease the $|\beta_T|$ estimate. But the $|\beta_T|$ estimate may be already overestimated as discussed earlier. These two factors may balance out to lead a pseudo-good estimation. This type of errors can be avoided by using variable (T_{oa}-T_{in}) instead of using T_{oa} or by correcting the model using a linear expression of T_{in} as a function of T_{oa}. For the Other day type, meaningful estimates are not available.

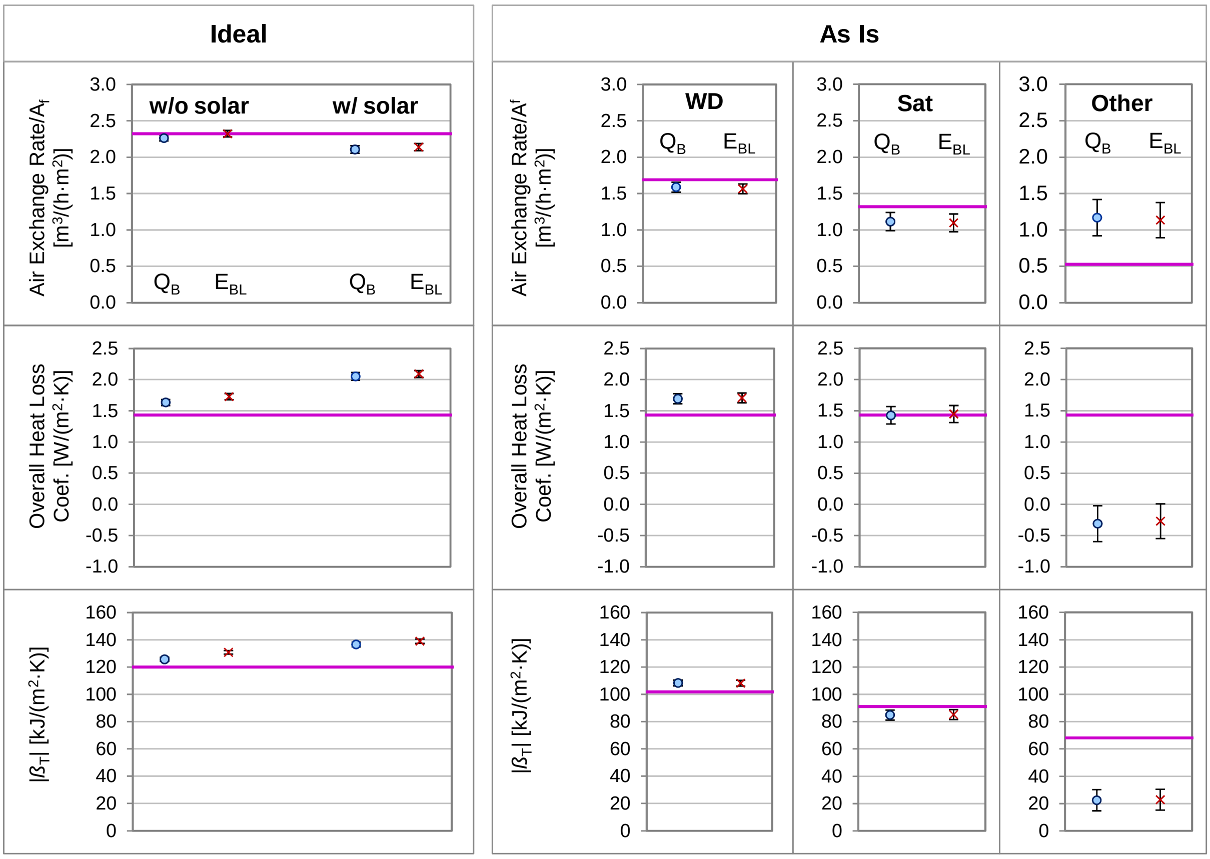

Fig. 7. Parameter estimates from synthetic daily data for the ideal and as-is cases. For each of three parameters, the assumed true value is shown as a solid line, and the parameter estimates using Q_B and E_{BL} are shown as a circle and a cross, respectively, along with the standard errors shown as bars.

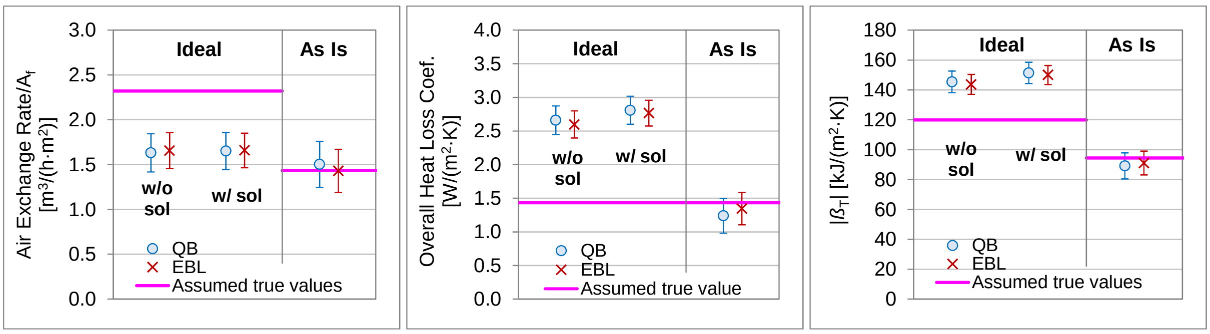

Fig. 8. Parameter estimates from synthetic monthly data for the ideal and as-is cases. For each of the parameters, the assumed true value is shown as a solid line, and the parameter estimates using Q_B and E_{BL} are shown as a circle and a cross, respectively, along with the standard errors shown as bars.

The parameter estimates using monthly interval data are presented in Fig. 8. In the ideal cases using monthly interval data, the estimated parameters have larger biases compared to the results from the daily interval data. This may be due to the large collinearity between the outside air temperature and humidity ratio variables in the monthly data, as the VIFs in **Error! Reference source not found.** shows large increase of the collinearity between T_{oa} and $X(W_{oa}\text{-}W_{in})$ in the monthly interval data. This indicates the model using monthly data is not able to separate effects of T_{oa} and $X(W_{oa}\text{-}W_{in})$ well. The reason for the good agreements between the estimates and the assumed true values in the as-is case using monthly data is not clear.

The T_{in}^* estimates for the synthetic data sets are shown in Fig. 9 along with the distribution of the daily average indoor air temperatures in the building. The physical meaning of this parameter changes with the structure of the models which mathematical expressions and approximate values can be found in Table 6. Both E_{BL} and Q_B models have good estimations for T_{in}^* in the ideal cases. In the as-is case, the bias increases as the unconditioned hours increases. The T_{in}^* estimated from E_{BL} models appear to be more stable over the different data sets, around a few degrees below the T_{in} when the HVAC systems are on for at least 16 hours per day. Based on these features of the T_{in}^* estimates from E_{BL} models, it is possible to create a rule of thumb for checking the estimated models. For example, the T_{in}^* should be in the range between the indoor air temperature and 2°C to 3°C below it; If the T_{in}^* is far away from the range, the model may not be reliable due to any possible reasons such as metering errors, model misspecifications, building operation changes during the data period, etc.

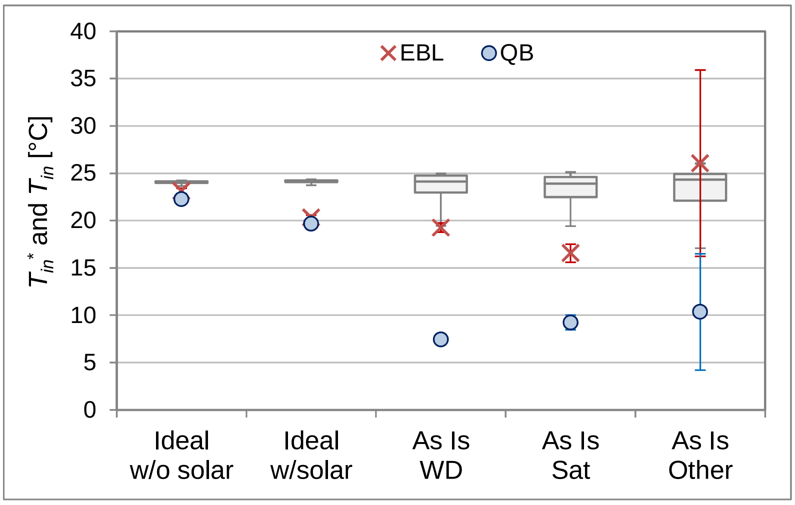

Fig. 9. Estimates of T_{in}^* and the distributions of T_{in}. For each case, estimates using E_{BL} and Q_B are shown with the standard errors, and the annual distribution of daily average T_{in} is presented by box and whisker plot.

Table 6. Physical meaning of the reference parameter T_{in}^* for different models used for synthetic data. Expected values are given.

Case	E_{BL}	Q_B
Ideal w/o sol	$T_{in} - \frac{Q_{occ}}{UA_s + m_v c_p} \sim 23.2°C$	$T_{in} \sim 24°C$
Ideal w/ sol	$T_{in} - \frac{Q_{occ} + a'_{sol}}{UA_s + m_v c_p + b_{sol}} \sim 21.8°C$	$T_{in} - \frac{a'_{sol}}{UA_s + m_v c_p + b_{sol}} \sim 22.6°C$
As is	$T_{in} - \frac{Q_{occ} + a'_{sol} - (1-f)E_E}{UA_s + m_v c_p + b_{sol}}$ $\sim 21.8°C$ if $f=1$	$T_{in} - \frac{Q_{occ} + a'_{sol} + fE_E}{UA_s + m_v c_p + b_{sol}}$ $\sim 13.8°C$ if $f=1$

4.2. Application to the Data from Actual Buildings

The air exchange rates estimated from the daily and monthly interval data from three dormitory buildings are compared to the measured values in Fig. 10, 16, and 17, and the VIFs of the variables are shown in Table 7. To see its effectiveness as a validation tool, the parameter T_{in}^* is also presented for each model. Overall, the estimates using daily data have similar values between Q_B and E_{BL} models for each building, but using monthly data, the estimates from E_{BL} models have better and stable results.

The proximity of the estimated parameters and the measured values vary between buildings. The HS building has the best estimations for the daily data, but it is underestimated for the monthly data, which is consistent with the results from the ideal cases of the simulation building. The T_{in}^* for the HS building falls in the expected range. The MF building has comparable results between daily and monthly data unlike other buildings. The VIFs of the monthly data for the building MF are small compared to the dataset for the other buildings, which should be due to lack of the data for July and August, the most hot and humid months. This less collinearity might be the reason for the similar results between the daily and monthly data.

Monthly data consist of small amount of data, and the estimates can be strongly influenced by anomalies. This seems to be the case with the building HB which appears to have changes in the outside air flow rate during the data period. The estimate from the monthly Q_B model for the HB building has about 140% higher than the measured value with a very low statistical significance. This seems to be caused by the collinearity between E_E and ($W_{oa}-W_{in}$), which can be seen in the high VIFs for these variables compared to the other datasets in Table 7. In fact, the effect of the E_E variable is overestimated around 5 times as the normal level. These abnormal estimates are alerted by the T_{in}^*; the estimate of T_{in}^* for the monthly Q_B model for the HB building is near 50°C which is not a realistic value based on the rule of thumb discussed earlier.

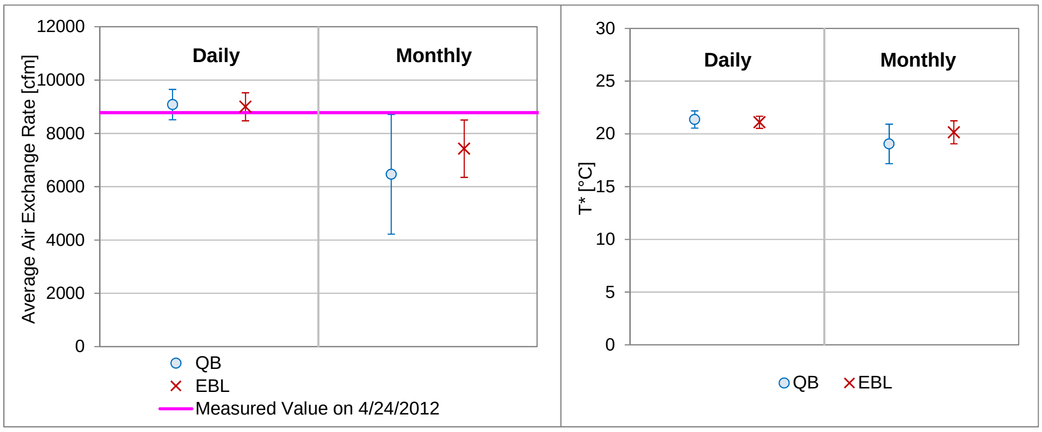

Fig. 10. Daily average outside air flow rate (left) and T_{in}^* (right) estimated for building HS comparing the estimates from daily and monthly interval data. Two different data periods are used. The standard error is shown with bars for each estimate. 1 cfm = 1.699 m^3/h.

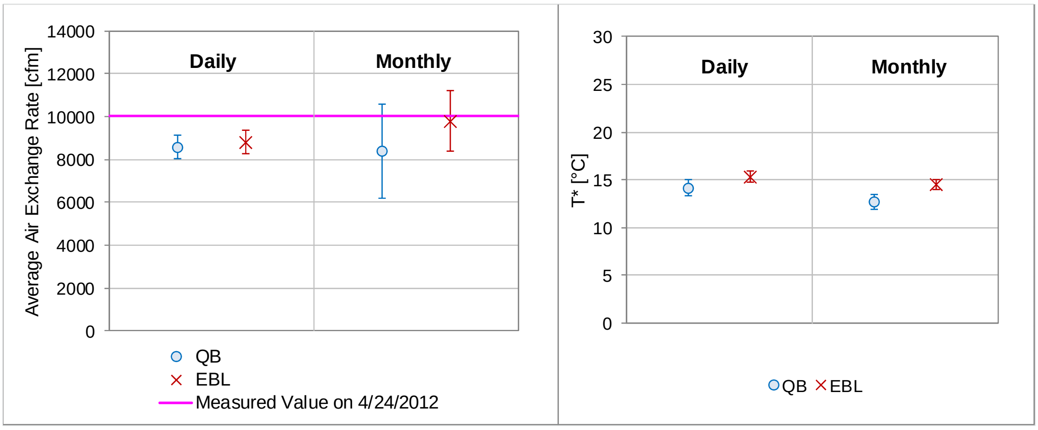

Fig. 11. Daily average outside air flow rate (left) and T_{in}^* (right) estimated for building MF comparing the estimates from daily and monthly interval data. The standard error is shown with bars for each estimate. 1 cfm = 1.699 m^3/h.

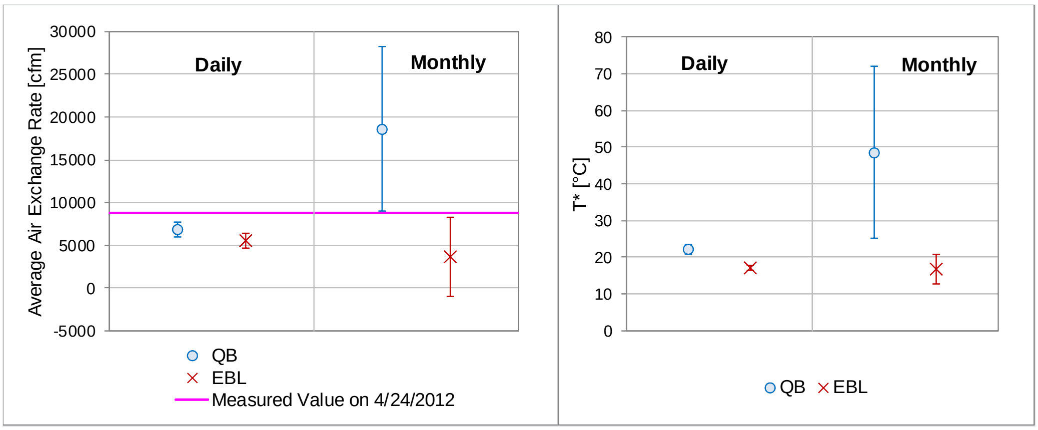

Fig. 12. Daily average outside air flow rate (left) and T_{in}^* (right) estimated for building HB comparing the estimates from daily and monthly interval data. The standard error is shown with bars for each estimate. 1 cfm = 1.699 m^3/h.

Table 7: VIFs for explanatory variables in the models for three dormitory buildings. For each set of variables, the values for daily and monthly data are compared.

Explanatory variable	HS				MF				HB			
	Daily	Monthly	Daily	Monthly	Daily	Monthly	Daily	Monthly	Daily	Monthly	Daily	Monthly
T_{oa}	2.95	14.53	2.86	4.25	2.17	5.23	2.13	2.48	2.72	11.39	2.66	4.01
$X(W_{oa}-W_{in})$	3.16	15.28	2.86	4.25	2.33	5.52	2.13	2.48	2.96	21.51	2.66	4.01
E_E	1.13	3.69			1.11	2.33			1.15	6.19		

5. Conclusions

The E_{BL} and Q_B models generally have a similar level of accuracy in the parameter estimation. However, the effects of the variables E_E and XE_E (i.e. k_s and k_l) in the Q_B models cannot be estimated properly in some cases and the inclusion of the variable may cause unexpected deviations in the parameter estimates, hence, Q_B models require more careful model selections compared to E_{BL} models. The estimations using daily data are fairly accurate when the HVAC systems are on for longer hours in the day. In the synthetic data for the commercial reference building model, meaningful estimates have been obtained for the schedules with the HVAC operation for 12 hours a day and longer (WD and Sat schedules). This indicates the method does not require a strict conformance to the constant parameter assumption, and the building without continuous HVAC operation can still be analyzed using this method by separating data into the day types with the same operation schedules. Meanwhile, the use of monthly data should be warned because of the large collinearity between the outside air temperature and humidity ratio and high sensitivity to the anomaly.

This method is applicable when the non-cooling electricity, cooling, and heating energy uses are metered separately. The method relies on the correct measurement; before the parameter estimation, one should check the validity of the data using appropriate techniques. It is often the case with the actual buildings that the building operations change during the modeling period. Such changes can be detected by analyzing the model residuals. The proposed reference parameter T_{in}^* may be used to detect some problems in the metered energy data and model misspecifications. The advantage of this parameter is the acceptable range is predictable without special knowledge of the building. The method to establish reasonable ranges for T_{in}^* under given conditions may be developed in the future study. The estimation of the outside air flow rate depends on the outdoor air humidity ratio variable, and if the data lacks hot and humid ambient conditions, the estimates may not be reliable. This can be caused by missing data but also resulted from the dry climate where the building stands. The applicability of the method to the different climate zones should be scrutinized.

6. Acknowledgements

Authors would like to thank *CC*® Team at Energy Systems Laboratory for the on-site measurements and useful information about the buildings. The monitored data used in this study was obtained from Texas A&M University Utilities & Energy Services Department, and we would like to thank all the personal for making the data available to us.

Nomenclature:

A_f = conditioned floor area of building, m^2
A_s = exterior surface area of building, m^2
c_p = specific heat at constant pressure, kJ/kg·K
E_{BL} = Energy balance loads, J
h_v = specific heat of vaporization, kJ/kg·K
k_l = ratio of internal latent loads to total internal sensible loads of building
k_s = multiplicative factor for converting Q_{LR} to total internal sensible loads
m_v = outside air exchange (ventilation) flow rate, kg/s
Q_B = Building thermal loads, J
E = metered energy use inside the building, J
T = dry-bulb temperature, °C
U = overall building envelope heat loss coefficient, $W/m^2 \cdot K$
W = humidity ratio, kg_w/kg_{da}
X = indicator variable
β = parameter of regression models

Subscripts:

E = whole building electricity
LE = whole building electricity (lights and equipment)
C = whole building cooling
H = whole building heating
oa = outside air
in = indoor air
sol = solar

References:

ASHRAE. (2009). *2009 ASHRAE Handbook: Fundamentals*. American Society of Heating, Refrigeration and Air-Conditioning Engineers, Atlanta, GA.

Baltazar, J. C., Sakurai, Y., Masuda, H., Feinauer, D., Liu, J., Ji, J., Claridge, D.E., Deng, S., and Bruner, H. (2007). *Experiences on the Implementation of the 'Energy Balance' Methodology as a Data Quality Control Tool: Application to the Building Energy Consumption of a Large University Campus*. Paper presented at the Seventh International Conference for Enhanced Building Operations, San Francisco, CA.

Baltazar, J., Claridge, D.E., Ji, J., Masuda, H., and Deng, S. (2012). *Use of First Law Energy Balance as a Screening Tool for Building Energy Data: Part II - Experiences of Its Implementation as a Data Quality Control Tool*. ASHRAE Transactions, 118, 167-174.

Claridge, D. E., Bensouda, N., Lee, S. U., Wei, G., Heinemeier, K., and Liu, M. (2003). *Manual of Procedures for Calibrating Simulations of Building Systems*, submitted to the California Energy Commission Public Interest Energy Research Program. (HPCBS#E5P23T2b). Retrieved July 1, 2012, from http://btus.lbl.gov/HPCBS/Element_5/02_E5_P2_3_2.html

Deng, S. (1997). *Development and Application of a Procedure to Estimate Overall Building and Ventilation Parameters from Monitored Commercial Building Energy Use*. (Master of Science, Texas A&M University).

DOE. (2011). *Commercial Reference Buildings*. Retrieved June 20, 2012, from http://www1.eere.energy.gov/buildings/commercial_initiative/reference_buildings.html

EnergyPlus (ver. 6). U.S. Department of Energy. http://apps1.eere.energy.gov/buildings/energyplus/

Flouquet, F. (1992). *Local Weather Correlations and Bias in Building Parameter Estimates from Energy-Signature Models*. Energy and Buildings, 19(2), 113-123

Haan, C. T. (2002). *Statistical Methods in Hydrology* (2nd. ed.) John Wiley and Sons.

Hammarsten, S. (1987). *A Critical Appraisal of Energy-Signature Models*. Applied Energy, 26(2), 97-110.

NCDC. (2012). QCLCD. Retrieved July 1, 2012, from http://cdo.ncdc.noaa.gov/qclcd/QCLCD

R Core Team. (2011). *R: A Language and Environment for Statistical Computing* (ver. 2.14.0). Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org

Rabl, A. (1998). *Parameter Estimation in Buildings: Methods for Dynamic Analysis of Measured Energy Use*. Journal of Solar Energy Engineering-Transactions of the ASME, 110, 52-66.

Rabl, A., and Rialher, A. (1992). *Energy Signature Models for Commercial Buildings - Test with Measured Data and Interpretation*. Energy and Buildings, 19(2), 143-154.

Reddy, T. A., Deng, S., and Claridge, D. E. (1999). *Development of an Inverse Method to Estimate Overall Building and Ventilation Parameters of Large Commercial Buildings*. Journal of Solar Energy Engineering, 121, 40-46.

Reddy, T. A., Kissock, J. K., Katipamula, S., and Claridge, D. E. (1994). *An Energy Delivery Efficiency Index to Evaluate Simultaneous Heating and Cooling Effects in Large Commercial Buildings*. Journal of Solar Energy Engineering, 116(2), 79-87.

Shao, X. (2006). *First Law Energy Balance as a Data Screening Tool*. (Master of Science, Texas A&M University).

Shao, X., and Claridge, D. (2006). *Use of first law energy balance as a screening tool for building energy data, part I - Methodology*. ASHRAE Transactions 2006, Vol. 112, Pt. 2, 717-731.