USE OF FIRST LAW ENERGY BALANCE AS A SCREENING TOOL FOR BUILDING ENERGY USE DATA: EXPERIENCES ON THE INCLUSION OF OUTSIDE AIR ENTHALPY VARIABLE

Hiroko Masuda
Graduate Research Assistant

Jing Ji, Ph.D.
Research Associate

Juan-Carlos Baltazar, Ph.D.
Associate Research Engineer

David E. Claridge, Ph.D., P.E.
Professor / Director

Energy Systems Laboratory, Texas Engineering Experiment Station
Texas A&M University, College Station, TX

ABSTRACT

Quality controlled energy-use data is the foundation of energy performance evaluation for a building. The "Energy Balance Load" (E_{BL}), a parameter derived from the first law of thermodynamics based on a whole-building energy analysis, has been theoretically proved to be an effective tool for verifying whole-building energy-use data (Shao and Claridge, 2006). Quality control methodology using E_{BL} has been proposed and applied to more than one hundred buildings on a large university campus by Baltazar et al. (2007). They picked the outside air dry-bulb temperature (T_{OA}) as the explanatory variable of E_{BL}, and used a plot of E_{BL} versus T_{OA}, called energy balance plot, to find faulty behavior in the data by visually observing the pattern. It has been demonstrated that this methodology can detect significant data problems caused by variety of reasons such as scale factor error and mislabeled meter successfully.

This paper presents a possible enhancement on the existent E_{BL} analysis technique by using the outside air enthalpy (h_{OA}) as the explanatory variable of E_{BL} instead of T_{OA}. This enthalpy based analysis accounts for the effect of latent load on E_{BL}, and therefore, may enhance the data screening capability for buildings operated at locations with hot and humid climate. Numerical threshold of data screening proposed by Masuda et al. (2008) has been applied to this enthalpy based methodology to determine the difference in the results of data screening between enthalpy based analysis and temperature based analysis.

BACKGROUND

Energy-use data provides valuable information that can be used for energy analysis to determine and improve building energy performance. And with recent advances in energy-use metering – increased functionality at lower costs – obtaining these data in a cost-effective manner is now becoming a standard practice (Sullivan, et al., 2007). Raw data of energy-use often includes misbehavior, and the demand for feasible method of data quality control to obtain usable data for energy analysis is emerging.

A simple but effective method of data quality control using Energy Balance Method has been proposed, and its application to more than one hundred buildings on the Texas A&M University campus has been illustrated by Baltazar, et al (2007). The methodology has successfully detected data problems such as scale factor error and mislabeled meter. This methodology is based on the known characteristic of Energy Balance Load (E_{BL}), a parameter evaluated based on the whole-building energy balance studied by Shao and Claridge (2006). E_{BL} has been represented as an overall linear relationship with the outside air dry-bulb temperature (T_{OA}) regardless of the type of the secondary HVAC system. Then the E_{BL} for a building evaluated from measured whole-building energy-use data of electricity, chilled water, and heating hot water forms a certain linear pattern when it is plotted versus outside air dry-bulb temperature. Knowing the pattern, misbehaved data can be detected visually.

To obtain consistent data screening results, independently of visual detection by data analysis experts, Masuda et al. (2008) have developed a technique to construct statistical control limits for one-year period E_{BL} data as a function of T_{OA}. E_{BL} regression model as a function of T_{OA} has non-constant variance due to the existence of latent load; the high temperature region has larger E_{BL} variability than the low temperature region has at hot and humid climate. To solve this problem, this technique enables the development of non-constant statistical bounds for a prescribed uncertainty level through overall temperature range based on the local variability of E_{BL} data. Alternatively, latent load effect can be included in E_{BL} analysis by using outside air enthalpy

(h_{OA}) as an explanatory variable. Ji et al. (2008) has presented the analytical study of E_{BL} as a function of h_{OA}, and have shown that using h_{OA} as an explanatory variable improves the capability of detecting faulty data at high outside temperature.

Following these two studies on enhancement to the data quality control using Energy Balance Method, this paper illustrates the application of E_{BL} as a function of h_{OA} to one-year period of actual energy use data for three buildings.

ENERGY BALANCE LOAD

The general derivation of the "Energy Balance Load" screening methodology comes from the first law of thermodynamics. The process is modeled as a semi-empirical methodology based on analytic redundancy (Shao and Claridge, 2006) applied to the whole building energy-use data. For a whole-building thermodynamic model, the heat flow rates and the rates of enthalpy flow across the boundary of its control volume and the rates of work performed on the building may be broken into its major components. The lumped form of the energy balance equation for a building can be expressed as:

$$\frac{d}{dt}\bar{E} = \bar{Q}vent + \bar{Q}solar + \bar{Q}cond + \bar{Q}occ + \bar{W}bheat - \bar{W}bcool + f\bar{W}bele \qquad (1)$$

Where E, is the energy storage in the building; *Wbele* is the whole building electricity use for lighting and equipment (non HVAC electric use); *Wbcool* is the whole building Chilled Water consumed to remove heat from the building; and Heating Hot Water required to provide heat in the building is represented by the term *Wbheat*; *Qsolar* is the solar radiation through the envelope; *Qvent* is the ventilation air and infiltration via doors, windows, or air-handling units; *Qcond* is the heat transmission through the building structure; and *Qocc* is the heat gain from occupants. The factor f is the portion of electricity that is converted to heat and appears as load within the building, there may be a time delay in this term relative to the actual time when the electricity is used. This equation is intended to capture the relevant features of the building energy-use without the complexity of the details such as the spatial variations of the temperatures inside and outside the building. Therefore, if the analysis is made on the basis of a period equal or greater than a day the equation can be considered quasi-steady. If it is arranged in a practical way, with the parameters that are typically metered and monitored in buildings, the equation could be represented as

$$\begin{aligned} E_{BL} &= \bar{W}bheat - \bar{W}bcool + f\bar{W}bele \\ &= -(\bar{Q}vent + \bar{Q}solar + \bar{Q}cond + \bar{Q}occ) \end{aligned} \qquad (2)$$

In this equation, the denominated "Energy Balance Load" (E_{BL}) term, represents a relationship between the metered parameters in the energy analysis. Shao and Claridge (2006) have proved that the E_{BL} parameter is independent of the type of air handling unit that is used in the building HVAC system. A typical parametric representation of the E_{BL} parameter as a function of the outside temperature follows a predominant line behavior, as shown in Figure 3. A more detailed parametric study can be found in Shao (2005). The values of the E_{BL} parameter are influenced by uncertainties of the instruments used for measurement of the energy-use and the incomplete model used for its formulation.

QUALITY CONTROL USING ENERGY BALANCE METHODLOGY

Knowing the mean structure of the Energy Balance Load as a function of the outside air temperature, it is possible to assemble a procedure to verify the energy-use data in a building is appropriate, provided that the electricity, chilled water and heating hot water are measured.

Figure 1 shows a typical data screening carpet plot used for energy data quality control by Baltazar, et al. (2007). The carpet plot includes four charts: an energy balance plot, which is a scatter plot of E_{BL} vs. T_{OA}, the corresponding time series plot of E_{BL}, a scatter plot of respective consumption data for electricity, chilled water, and heating hot water versus outside air dry-bulb temperature, and their corresponding time series plot. In the data screening process, data analysis expert first look at the energy balance plot to find outliers or pattern misbehavior. If any unusual E_{BL} data points are found, the other plots are referred to identify the energy-use data that caused it. Then the causality is assessed to determine if the data needs correction.

Control limits for the energy balance plot, which is an extension of statistical threshold for the data screening, has been proposed by Masuda et al. (2008). Energy Balance Load as a function of T_{OA} loses its linearity in the high temperature region (Shao and Claridge 2006). This is due to the large latent load in the high outside temperature season under hot and humid climate. Since the influence of latent load is large, the variability in the E_{BL} regression model with T_{OA} as the explanatory variable increases in the high temperature region. To construct control limits with consistent uncertainty level for all over the temperature, this method estimates a variable function ($S_{\hat{E}_{BL}}{}^2$) for the E_{BL} regression model, which is the variable variance as a function of T_{OA}, from the local variances. Upper control limit (UCL), center line (CL) and lower control limit (LCL) as a function of T_{OA} are defined as following equations.

$$\begin{aligned} UCL(T_{\mathrm{OA}}) &= \hat{E}_{BL}(T_{\mathrm{OA}}) + kS_{\hat{E}_{BL}}(T_{\mathrm{OA}}) \\ CL(T_{\mathrm{OA}}) &= \hat{E}_{\mathrm{BL}}(T_{\mathrm{OA}}) \\ LCL(T_{\mathrm{OA}}) &= \hat{E}_{BL}(T_{\mathrm{OA}}) - kS_{\hat{E}_{BL}}(T_{\mathrm{OA}}) \end{aligned} \tag{3}$$

where

$\hat{E}_{BL}$ = E_{BL} regression on the mean, change-point regression model is allowed to express non-linearity of E_{BL} as a function of T_{OA}

$S_{\hat{E}_{BL}}$ = square root of variance function, corresponds to estimated standard uncertainty

k = coverage factor, a multiplicative number to define the distance of the limits from the center line in terms of $S_{\hat{E}_{BL}}$

Note that when the model residuals don't have dependency on T_{OA}, $S_{\hat{E}_{BL}}$ equals the prediction error on inference of an individual response, and UCL and LCL correspond to the prediction interval. If a newly observed E_{BL} falls out of the region bounded by UCL and LCL, the E_{BL} point may includes misbehaved energy-use data, and causality analysis will be performed. Since the algorithm can be employed in computer programs, it is possible to develop automated E_{BL} data screening process for large-volume data processing and for consistent screening results.

USING ENTHALPY FOR ENERGY BALANCE ANALYSIS

The problems related to latent load described in the previous part is aroused because dry-bulb temperature alone is used as the explanatory variable of E_{BL} model. The possible solution for this is to include humidity variable into the model.

The enthalpy of moist air is the sum of the enthalpy of the dry air and of the water vapor comprising the mixture. In terms of temperature and humidity ratio, the enthalpy of moist air is expressed as following equation (Kreider et al., 2005).

$$h = c_{pa}T_d + W(h_{g,ref} + c_{pw}T_d) \tag{4}$$

where

h = specific enthalpy of most air, Btu/lb$_{\mathrm{da}}$

c_{pa} = specific heat of dry air, Btu/(lb$_{\mathrm{da}}$-°F)

c_{pw} = specific heat of water vapor, Btu/(lb$_{\mathrm{w}}$-°F)

T_d = dry-bulb temperature, °F

$h_{g,ref}$ = enthalpy of saturated water vapor at reference temperature of 0 °F, Btu/lb$_{\mathrm{w}}$

W = humidity ratio, lb$_{\mathrm{w}}$/lb$_{\mathrm{da}}$

For outdoor air temperature range, c_{pa} and c_{pw} are assumed to be constant, and the enthalpy is just a function of dry-bulb temperature and humidity ratio. Then there is an advantage of using enthalpy as an explanatory variable since sensible and latent load of outside air can be represented by a single variable. It allows us to analyze E_{BL} by simple linear regression and to visualize the relation in a two-dimensional plot as same way as using T_{OA}.

Ji et al. (2008) has analyzed the structure of E_{BL} as a function of the outside air enthalpy (h_{OA}) based on the simplified air side load simulation, using the bin data of outside air temperature and wet-bulb temperature, for four secondary systems: Single-duct constant air volume with terminal reheat (CVRH), Dual-duct constant air volume (DDCV), Single-duct variable air volume (SDVAV) and Dual-duct variable air volume (DDVAV). The results are plotted as a function of T_{OA} in Figure 3 and as a function of h_{OA} in Figure 4. In both plots, the patterns are consistent regardless of the type of the secondary systems. And the result indicates E_{BL} as a function of h_{OA} has better linearity than that as a function of T_{OA} in the hot and humid outdoor condition.

Figure 2 is the E_{BL} data screening carpet plot using h_{OA} in the scatter plots for the same data set as used in Figure 1. In regression analysis, smaller variance means that the explanatory variable accounts for the response variable better. For this data set, the variance in the region under hot and humid outdoor condition is smaller in h_{OA} plot in Figure 2 than in T_{OA} plot in Figure 1. RMSE of the four-parameter change-point (4P-CP) regression model (Kissock, et al. 2002) for E_{BL} versus T_{OA} is 104.3 [Btu/day-ft^2] and for E_{BL} versus h_{OA} is 72.0 [Btu/day-ft^2], it has been improved by 31% by using h_{OA} as the explanatory variable.

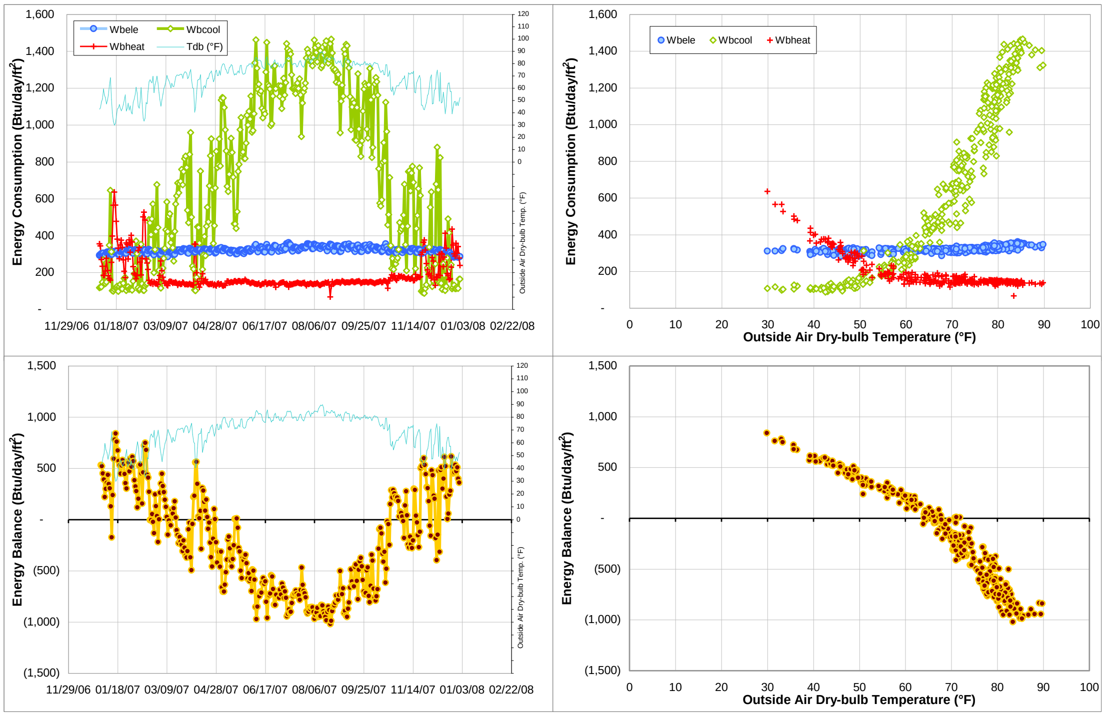

Figure 1 ‘Energy Balance’ carpet plot for data screening for an office and lab building based on the energy-use data during 1/1/2007 – 12/31/2007. Outside air dry-bulb temperature is used in the scatter plots.

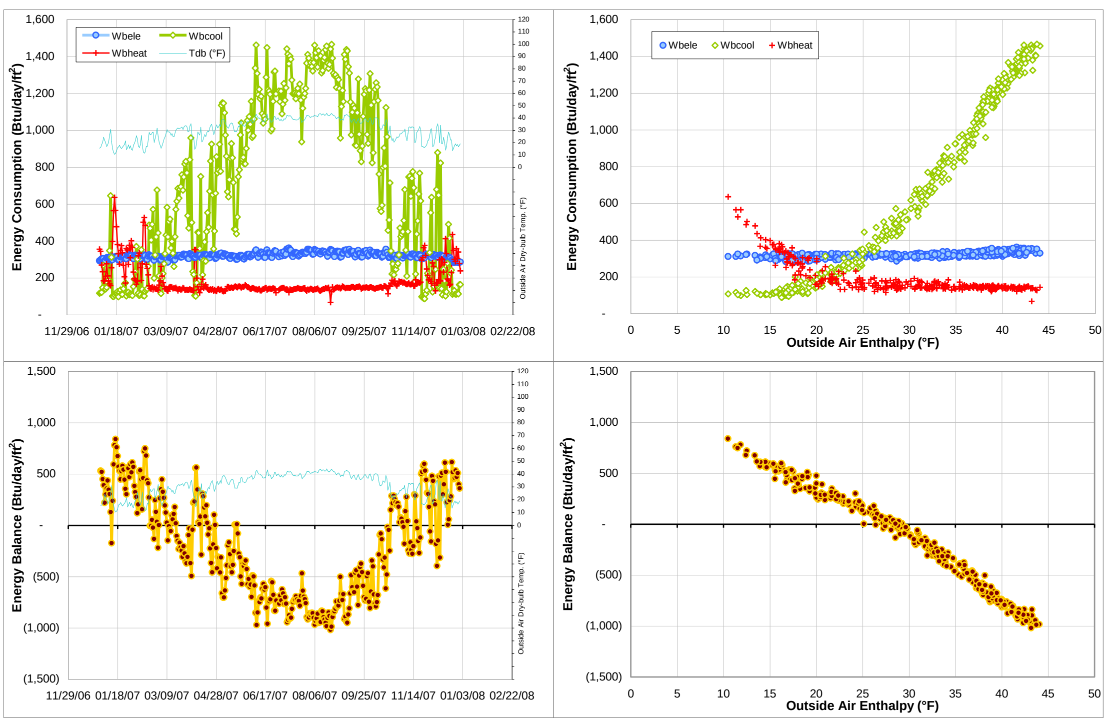

Figure 2 ‘Energy Balance’ carpet plot for data screening for an office and lab building based on the energy-use data during 1/1/2007 – 12/31/2007. Outside air enthalpy is used in the scatter plots

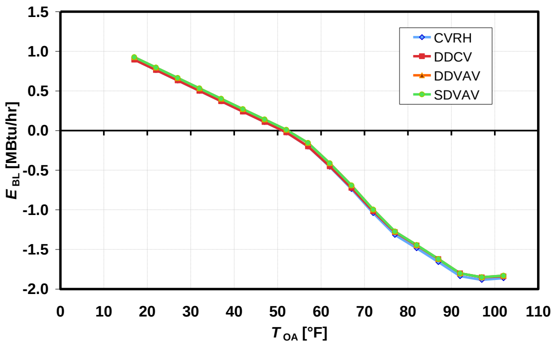

Figure 3 E_{BL} as a function of T_{OA} based on the simplified air side load simulation

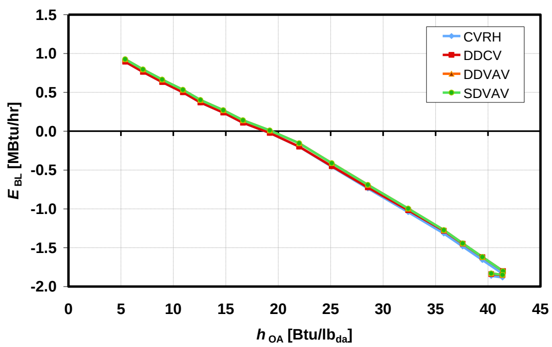

Figure 4 E_{BL} as a function of h_{OA} based on the simplified air side load simulation

CASES OF STUDY

Control limits of E_{BL} as a function of T_{OA} and as a function of h_{OA} have been developed for sets of daily data for one year from three buildings on Texas A&M University campus. Each case has two plots, one for T_{OA} and the other for h_{OA}. In each plot, the center line and the two levels of control limits for k=2 and k=3 are plotted in conjunction with the E_{BL}. The four-parameter change-point (4P-CP) regression model was used for development of the control limits. The comparison of RMSEs is given for each building using T_{OA} and using h_{OA} as the explanatory variable in the model.

For all buildings, the daily energy-use for electricity, chilled water and heating hot water was totaled from measured whole-building hourly data, and daily E_{BL} parameter was evaluated using the Eq. (2). The daily average outside air dry-bulb temperature and the daily average outside air enthalpy were calculated from the hourly observation of quality controlled local climatological data (QCLCD) for College Station, TX, acquired from National Climatic Data Center.

Case I: Office Building

This office building has an area of 65,688 ft^2. The period of the E_{BL} data is from 7/1/2005 through 6/30/2006; three data out of 365 days were excluded as outliers that have influence on the model. The model variance of this building moderately increases with T_{OA} under hot and humid season as shown in Figure 5. Meanwhile, if h_{OA} is used as the explanatory variable as in Figure 6, the model variance appears to be uniform over all the enthalpy range, and the distance of the UCL and LCL under hot and humid season decreases significantly. The RMSE of the model decreases by 43 %.

Table 1 Comparison of RMSE for Case I

	Explanatory Variable	
	T_{OA}	h_{OA}
RMSE [Btu/day-ft^2]	54.6	31.0

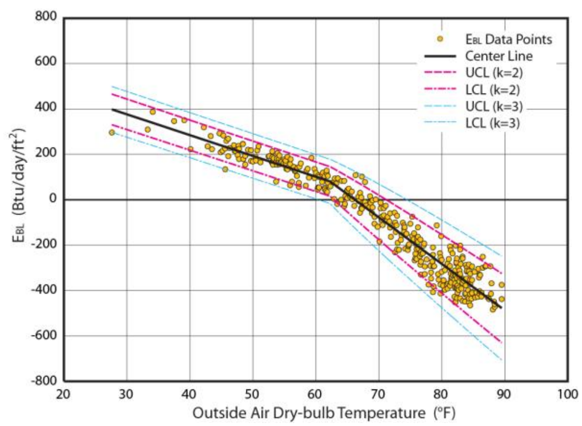

Figure 5 E_{BL} control limits as a function of T_{OA} in conjunction with the E_{BL} data for Case I

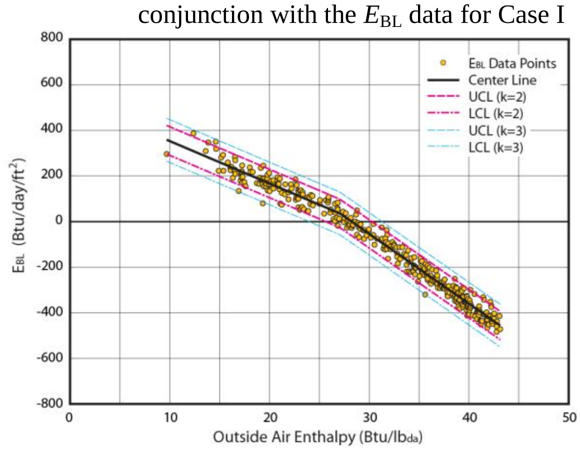

Figure 6 E_{BL} control limits as a function of h_{OA} in conjunction with the E_{BL} data for Case I

Case II: Office and Lab Building

This office and laboratory building has an area of 62,273 ft^2. The period of the E_{BL} data is from 6/1/2005 through 5/31/2006; one data out of 365 days were excluded as an outlier that has influence on the model. The model variance of this building strongly increases with T_{OA} under hot and humid season as shown in Figure 7. Presumably, having biological labs, this building requires high ventilation rate, which may leads to the strong dependence of E_{BL} on the outdoor humidity. Similarly to the Case I, using h_{OA} as the explanatory variable significantly decreases the distance of UCL and LCL under hot and humid season as shown in Figure 8. Again, the model variance appears to be uniform over all the enthalpy range. The RMSE of the model decreases by 33 %.

Table 2 Comparison of RMSE for Case II

	Explanatory Variable	
	T_{OA}	h_{OA}
RMSE [Btu/day-ft^2]	44.5	29.8

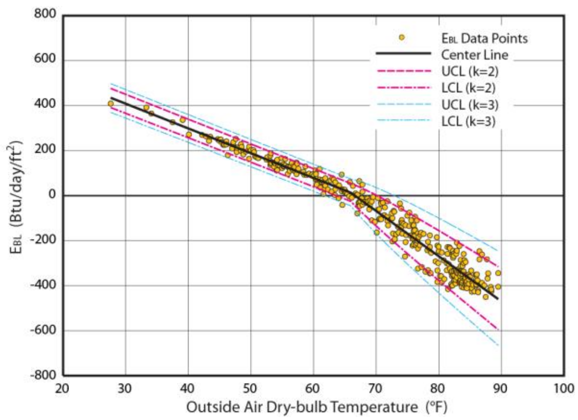

Figure 7 E_{BL} control limits as a function of T_{OA} in conjunction with the E_{BL} data for Case II

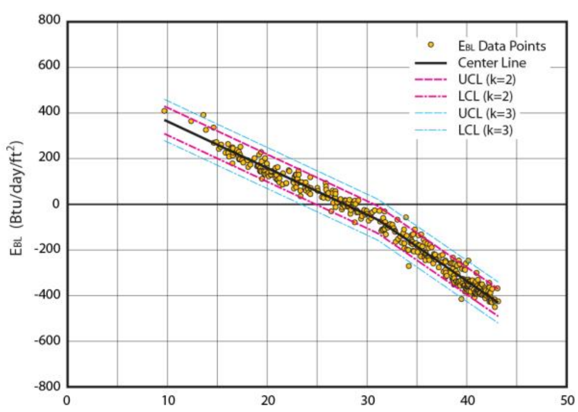

Figure 8 E_{BL} control limits as a function of h_{OA} in conjunction with the E_{BL} data for Case II

Case III: Residence Building

The Case III building is a dormitory and the area of the building is 59,541 ft^2. The period of the E_{BL} data is from 11/1/2006 through 10/31/2007; one data out of 365 days were excluded as an outlier that has influence on the model. The model variance of this building doesn't have remarkable dependence on T_{OA} as shown in Figure 9. Unlike the former two cases, using h_{OA} as the explanatory variable as shown in Figure 10 doesn't improve the model. In fact, this dormitory is usually vacant for a few months during summer, and the outside air intake might have been decreased during summer. This might be the reason for that the E_{BL} behavior doesn't indicate any influence by latent load during hot and humid season.

Table 3 Comparison of RMSE for Case III

	Explanatory Variable	
	T_{OA}	h_{OA}
RMSE [Btu/day-ft^2]	21.3	24.4

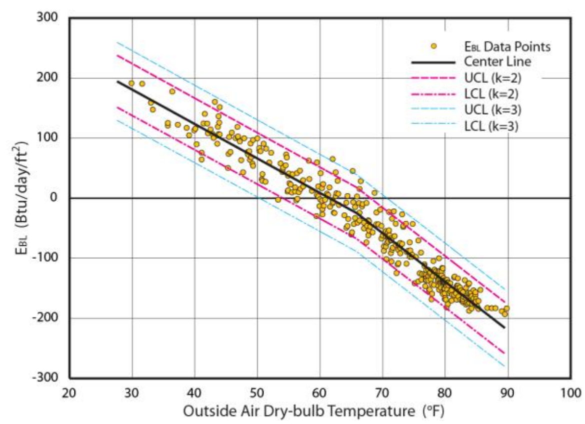

Figure 9 E_{BL} control limits as a function of T_{OA} in conjunction with the E_{BL} data for Case III

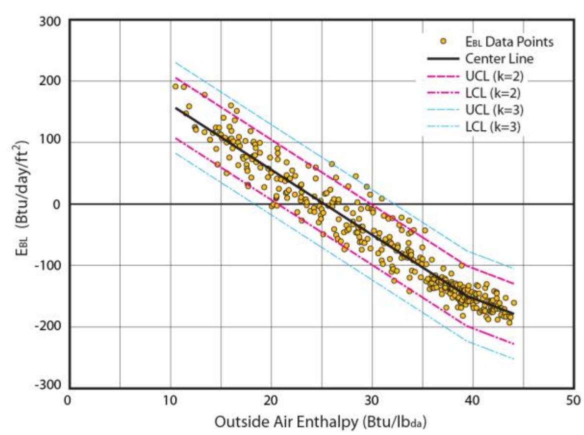

Figure 10 E_{BL} control limits as a function of h_{OA} in conjunction with the E_{BL} data for Case III

For such operation change, it is desirable to group the data based on known operational changes, and analyze those separately. For this case, the RMSE of the model using h_{OA} increases by 14.6 % from that of the model using T_{OA}.

CONCLUSIONS

Application of outside air enthalpy as the explanatory variable in E_{BL} regression model as a data screening tool was presented. For the case studies, when the dry-bulb model has increasing variance with the temperature, the enthalpy model shows better fitting and narrower bounds of control limits in the high temperature and humidity region. Then enthalpy model provides better capability of data screening for these cases. The other advantage observed in the case studies is that the enthalpy model appears to have uniform variance all over the enthalpy region. This may allow us to utilize well-established regression analysis techniques to detect outliers and leverage and to test model parameters and fitting, most of which based on the assumption of uniform variance.

However, enthalpy is a property which cannot be directly measured by physical sensors, and evaluation of enthalpy requires measurements of other properties in addition to dry-bulb temperature. Temperature model still has an advantage over enthalpy model when sufficient resources for additional measurement, time and skilled technicians are not available. Temperature model is also valuable for users because dry-bulb temperature is associated with physical sense, stronger than any other air properties, and the results in terms of temperature can be interpreted more easily than in terms of enthalpy. If the two models are used in parallel, it is necessary to study consistency of the data screening results by the temperature model and by the enthalpy model.

Since temperature models and enthalpy models were compared for only three buildings, the application for more buildings should be undertaken.

REFERENCES

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-use of a Large University Campus, *International Conference for Enhanced Building Operations*, San Francisco, CA. Nov.

Ji, J., Baltazar, J. C. and Claridge, D. E., 2008, Study of the Outside Air Enthalpy Effects in the Screening of Metered Building Energy Data, *The 16th Symposium on Improving Building Systems in Hot and Humid Climates*, Plano, TX. Dec.

Kissock, J. K., Haberl, J. S. and Claridge, D. E., 2002, *Development of a Toolkit for Calculating Linear, Change-point Linear and Multiple-Linear Inverse Building Energy Analysis Models*, ASHRAE Research Project RP-1050. Final Report, Nov.

Kreider, J. F., Curtiss, P. S. and Rabl, A., 2005, *Heating and Cooling of Buildings: Design for Efficiency*, 2nd ed., Kreider and Associates, Boulder, CO.

Masuda, H., Baltazar, J. C., Ji, J. and Claridge, D. E., 2008, Development of Data Quality Control Limits for Data Screening through the “Energy Balance” Method, *The 16th Symposium on Improving Building Systems in Hot and Humid Climates*, Plano, TX. Dec.

National Climatic Data Center, 2009, Quality Controlled Local Climatological Data: QCLCD (v.2.2), Retrieved from

http://cdo.ncdc.noaa.gov/qclcd/QCLCD

Shao, X. and Claridge, D.E., 2006, Use of First Law Energy Balance as a Screening Tool for Building Energy Data, Part I – Methodology, *ASHRAE Transactions*, Vol. 112, Part 2, QC-06-068.

Shao, X., 2005, *First Law Energy Balance as a Data Screening Tool*, M. S. Thesis, Mechanical Engineering Department, Texas A&M University, College Station, TX, May.