CASE STUDY OF APPLYING DIFFERENT ENERGY USE MODELING METHODS TO AN EXISTING BUILDING

Bin Yan, Ali M. Malkawi, and Yun Kyu Yi T.C. Chan Center for Building Simulation and Energy Studies University of Pennsylvania, Philadelphia PA 19104, United States

ABSTRACT

Various building energy use modeling methods have been applied to existing buildings in order to understand building energy performance and improve energy efficiency. There are widely used models based on physical principles and historical data. This study has used temperature-based regression, artificial neural network and EnergyPlus models to predict energy use of a laboratory building. The paper discusses the accuracy of different methods when predicting short-term and long-term whole building energy use. It also discusses the feasibility and limitation of analyzing component level energy use and evaluating energy savings potential.

INTRODUCTION

Various building energy use modeling methods have been applied to existing buildings in order to understand building energy performance and improve energy efficiency. For example, simplified models based on heat balance equations and detailed simulation such as EnergyPlus, DeST and DOE-2 can serve to optimize operating strategies (Liu and Claridge, 1998; Yan et al., 2009). Temperature-based regression models have been widely used for determining weather-adjusted savings (Kissock et al., 1998). Artificial neural network (ANN) models have been used to estimate energy savings due to heating, cooling, and lighting equipment retrofits (Cohen and Krarti, 1995).

These energy use models can be categorized into two groups, historical data based and physical principle based. Artificial neural network and temperaturebased regression models are developed from historical data. When developing physical principle based models, for example, using EnergyPlus, metered data are not required, though metered data might be necessary for model calibration.

Whole building energy use prediction is required when deriving annual consumption based on incomplete metered data. Analysis of component level energy use is important to understand building energy performance. Estimating energy savings of implemented retrofits and new strategies is essential for cost-benefit analysis of commissioning and retrofit projects. Energy use modeling tools can assist these tasks. The roles of different energy use modeling tools in these tasks should be explored. As shown in Figure 1, this study has used temperaturebased regression, artificial neural network and detailed simulation to predict whole-building energy use of a laboratory building. It compares the accuracy of short-term and long-term prediction among all three models. The paper also discusses the feasibility and limitation of different models when analyzing component level energy use and evaluating energy savings potential. It illustrates examples of using detailed simulation for component-level energy use analysis and estimation of energy savings.



Figure 1 Illustration of the case study content

This case study is based on a five-floor laboratory building with a total area of 11400 m². Three primary air handling units with heat recovery systems serve labs, offices, meeting rooms and common areas. Variable air volume (VAV) boxes with terminal reheat serve as terminal units in this building, along with a small proportion of constant air volume (CAV) boxes. In the air handling units, the air temperature set point after preheat coils is 10°C, and the supply air temperature setpoint is 11.1°C. The heat recovery system is enabled when outside air temperature is below 10°C. Steam will be supplemented through heat exchangers if the setpoint still cannot be maintained by heat recovery. Pre-cool through heat recovery is enabled when outside air temperature is above 28.3°C. In this laboratory building, chilled water from the plant, purchased steam and electricity are the three types of energy resources. Chilled water heat flow and steam mass flow are sub-metered. Approximately one-year data of an interval as short as five minutes are available.

However, due to communication loss and after ruling out the problematic data caused by sensor error, there are 3393 data points of hourly chilled water use and 6444 data points of steam use left for this study.

MODELING METHODS

Temperature-based Regression

Regression based on outside air temperature is a relatively easy-to-implement approach for energy use modeling. In commercial buildings, there is a strong linear relationship between energy use and outside air dry-bulb temperature.

Including other environment variables, such as wind speed and solar radiation, might improve the accuracy of a regression model. However, the degree of improvement might not be worth the complexity introduced by these additional variables. Moreover, these additional variables are not as easy to collect as temperature. Previous studies show that outside air dry-bulb temperature and dew-point temperature can explain more than 90% of the variation in energy use. It was reported that regression models based on outside air dry-bulb temperature could describe commercial building cooling and heating energy use with RMSEs of about 15 percent of the mean energy consumption (Kissock et al., 1998).

Figure 2(a) and (b) plot the normalized chilled water use and steam use according to the outside air drybulb temperature. The overall chilled water use and steam use show a piecewise linear relationship with outside air dry-bulb temperature. Therefore, this study uses piecewise linear regression models based on outside air dry-bulb temperature to describe chilled water use and steam use. The change points are determined through a grid search method using RMSE as the criterion. Humidity is included in the cooling model since there is significant latent ventilation load, as shown in Figure 2(a). It is an interesting phenomenon that chilled water use is negative sometimes when outside air temperature is low. This happens when the supply water temperature is even slightly higher than the air temperature. The water is cooled by the air and the chilled water use becomes negative. The negative chilled water use reveals a deficiency of the system.

Artificial Neural Network

Artificial neural network (ANN) can model nonlinear process based on historical data. ANN has been used to predict building energy use approximately since the 1990s. For the purpose of modeling cooling and heating energy use, widelyused inputs are weather variables including temperature, humidity, wind speed and solar radiation, and time-related variables. The reported error rates of short-term prediction (1h to 24h) can be as low as 1%-5%. Long-term prediction accuracies are also promising (Dodier and Henze, 2004).

In this study, data samples are on an hourly basis. The targets are chilled water use and steam use. The inputs include outside air dry-bulb temperature, humidity ratio, the hour of the day and the day of the week. We use

$$sin\left(\frac{2\pi N}{T}\right)$$
 (1)

$$\cos\left(\frac{2\pi N}{T}\right)$$
 (2)

together to represent the hour of the day and the day of the week (Dodier and Henze, 2004). When representing the hour of the day, N is the hour of the day and T = 24. Likewise, when representing the day of the week, N is the day of the week and T = 7. The training of neural network is implemented through Matlab (version R2010a) Neural Network Toolbox. In the model, there is one hidden layer with 15 neurons. The activation equation in the hidden layer is sigmoid, and linear in the output layer.



Figure 2 Chilled water and steam use versus outside air dry-bulb temperature

Detailed Physical Principles Based Models

Energy use models based on physical principles are well-known and well-studied for energy use prediction. Many simulation tools play an important role in the design for new buildings, such as Energy Plus, which is well developed.

Unlike models based on regression and artificial neural network, metered data is not required in building a model when using detailed simulation. However, in order to ensure accuracy, metered data is necessary in order to calibrate the model. Large amounts of inputs are required for the detailed simulation. In this study, EnergyPlus is used for energy consumption modeling. Metered electricity data is used to calibrate the lighting and plug-load settings. Monthly metered chilled water use and steam use are used to calibrate the model.

RESULT ANALYSIS

In this section, we analyze different applications of three energy use modeling methods to existing buildings. The applications discussed here include predicting whole building energy use, analyzing component level energy use and estimating energy savings. Annual energy use of the whole building is an indicator of energy performance. Analysis of component level energy use can help further understand building energy performance and find elements that most affect building energy use, where significant energy savings will come from. Estimating energy savings is essential for costbenefit analysis in commissioning and retrofit projects.

Predicting Whole Building Energy Use

Metered data are seldom complete. Sometimes, engineers need to estimate annual energy

consumption when there is only data for several months.

Artificial neural network and temperature-based regression are two widely used methods to predict energy use based on historical data. The coefficients in the neural network and the regression model are determined using 240 consecutive hourly data points as a testing set and the remaining data for training. A leave-one-out cross validation has been performed, considering 240 consecutive hourly data points as one sample. Then using all available data as training data, a one-year prediction has been made by these two models.

The accuracy is evaluated according to the coefficient of variation of the root mean square deviation,

$$CV = \frac{\sqrt{1/N\sum_{t}(y_{t} - \hat{y}_{t})^{2}}}{1/N\sum_{t}y_{t}}$$
(3)

Here, N is the number of test samples. In the comparison on an hourly basis, y_t is an hourly target value and \hat{y}_t is an hourly prediction. In the comparison on a daily basis, y_t is the accumulative metered consumption of 24 hours, and \hat{y}_t is the corresponding prediction.

Figure 3 shows a sample of an hourly prediction of ANN and regression models. Using hour of the day and day of a week as input features, the ANN model is able to learn from the variation due to occupancy. Therefore, compared with regression results, ANN is likely to have better performance on a short data time scale. As shown in Figure 3, the hourly regression results are consistent with the overall trend of metered data, while ANN model successfully predicted the spikes in energy consumption. As shown in Table 1, on an hourly basis, ANN model has a relatively higher accuracy.



Figure 3 Hourly prediction of ANN and Regression models

 Table 1

 Coefficient of variation of three modeling methods for hourly and daily whole building energy use prediction

	ANN		Regression		Detailed Simulation	
	Hourly	Daily	Hourly	Daily	Hourly	Daily
Chilled Water	25.03%	18.14%	27.31%	18.11%	32.80%	24.32%
Steam	14.48%	8.14%	16.44%	9.94%	23.00%	17.00%

However, if selecting a time scale equal to or longer than 24 hours, the effect of variation from factors such as internal load and solar radiation can be removed. Thus, on a daily or a longer time data scale, the ANN model does not show significant advantage over the regression model from the perspective of accuracy. Figure 4 shows that monthly prediction of ANN and regression are very close. The differences of one-year consumption for both chilled water and steam between ANN and regression are within 4%.

When applying detailed simulation to whole-building energy use prediction, approximately 5 months of chilled water use and 9 months of steam use are used for calibration to ensure accuracy. The model has been calibrated by monthly data, with coefficients of variation of 12.8% for monthly chilled water use and 12.2% for monthly steam use.

kWh/m²

Figure 4 shows that the monthly simulation prediction results of three modeling methods are similar. However, as shown in Table 1, the accuracies of hourly and daily prediction of detailed simulation are noticeably lower than ANN and regression models. Figure 5 plots the simulated results of detailed simulation versus measured data on a daily basis. In the plot of chilled water use, the samples are scattered uniformly around the diagonal line. While in the plot of steam use, simulation severely underestimates the steam use for a certain number of days, when daily average outside air temperature is between 0 °C to 10 °C. The underestimation might be caused by poor estimation of occupancy for a certain period. Malfunction of the preheating system might also exist in the operation and cause the deviation.



■ ANN ■ Regression ■ Simulation









Although the detailed building simulation model can be calibrated until its results closely match the measured data even on an hourly basis, it does not mean that the simulation model completely reflects the real system performance. There are uncertainties in using simulation models to mimic system performance of existing buildings. They come from lack of information for the inputs and limitations of the model (Haves et al., 2001). In most cases, inputs for plug loads, lighting and occupancy are impossible to measure accurately and the settings in the model are an estimation. Moreover, simulation models assume idealized behaviour of systems. Current models do not reflect imperfect and faulty mechanical operation in reality.

Due to the uncertainties of detailed simulation and the amount of efforts it requires, it is not the first choice to derive annual energy use of existing systems based on historical data. However, in some aspects, where metered data is not available, detailed simulation can provide adequate predictions, which we will discuss in the following sections.

Analyzing Component Level Energy Use

Understanding how much and where the energy has been consumed in a building can help discover energy savings potential. In this case study, the regression model estimates that annual chilled water use is 717 kWh/m², and the steam consumption is 673 kWh/m². The energy intensity is high, which makes it more important to figure out where the energy is consumed.

The metered data used for training of historical data based models is whole-building energy consumption. As a result, no component level energy use can be derived from the historical data based models, while simulation models based on physical principles are able to give details on component energy use. In this study, EnergyPlus is used to analyze component level energy use.

Figure 6 shows the components that account for heating and cooling. The results are plotted by outside air temperature, so that the effect of outside air temperature on component energy use can be easily visualized. Heat balance relationship changes according to outside air temperature. When outside air temperature is higher than indoor temperature, chilled water is mainly used to offset the ventilation load. Due to the large air exchange rate required by standards for laboratories, reheat is still necessary, which counteracts part of cooling provided by chilled water. In comparison, the amount of heat gain through building envelopes and solar radiation is insignificant. When the outside air temperature is lower than 10 °C, outside air provides free cooling. As outside air temperature decreases, steam use increases for preheating. In addition, there is more VAV terminal reheat for heat loss through envelopes, although the amount is insignificant. Compared with steam use for reheat and chilled water use due to high ventilation rate, other components have relatively less impact on energy use. Therefore, it is a priority to investigate the possibility of reducing air exchange rate for energy savings.



Figure 6 Component level heating and cooling analysis

Estimating Energy Savings

In order to evaluate energy savings of implemented retrofits, energy consumption before and after the retrofits is required. It is difficult to acquire a complete data set for one year before the retrofits and another after the retrofits. Moreover, the effect of weather on energy consumption needs to be removed in the comparison. Therefore, it is often necessary to generate two complete data sets from the available data with the same weather conditions. Then the problem turns into long-term whole building energy use prediction. Historical data based models can be applied as described in the previous section.

For decision-making, it is often necessary to estimate energy savings potential of new strategies before carrying them out. Without any data after changes are implemented, historical data based models cannot be used. Detailed simulation models based on physical principles can serve this purpose.

A simple example from this case study is to estimate energy savings potential for reducing the mechanical ventilation rate. The National Fire Protection Association recommended that occupied laboratories often have ventilation rates on the order of 8 to 10 ACH, but it could be as low as 4 ACH when unoccupied (NFPA, 2000). The building in this case study maintains a ventilation rate 8 to 10 ACH all the time. Another simulation has been performed, in which the minimum flow rate to laboratories is 4 ACH from 12AM to 7AM. 10.1% chilled water use and 16.7% steam use can be saved from this strategy.

DISCUSSION

For historical data based models, data quality and quantity are crucial to model accuracy. For instance, if there is only metered data in summer, artificial neural network and regression models will not be able to predict energy use in winter accurately since there is no information when outside air temperature is low. Besides, outliers in data will also affect neural network training and regression model. In this study, robust regression is used to reduce the weight of outliers in temperature based regression models. For further research, it might be necessary to investigate the criteria of data quality and quantity to ensure certain accuracy of historical data based models.

In this case, the ANN models do not show significant advantages over simple regression models. One possible reason is that since it is a primary system, the energy use has a very strong linear relationship with outside air temperature and humidity. Besides, the ANN models in this study only use two more features than the regression models, the hour of a day and the day of a week, in order to include the occupancy in the model. If data of more relevant features such as solar radiation and hourly electricity consumption are available, accuracy of ANN models might improve. Detailed physical principle based simulation provides more information about building performance than historical data based models. The accuracy is acceptable from an engineering perspective, although lower than historical data based models in this case study. A more careful calibration will improve accuracy. For example, it will be helpful to conduct a more detailed survey on the building and its system. Thus, the inputs of detailed simulation will be closer to actual conditions. It might also be helpful to do a daily or even hourly calibration, instead of merely calibrating by month. However, the time and efforts required are usually unaffordable in engineering projects. Moreover, it is difficult to eliminate uncertainties in the modeling. For example, the simulation model assumes idealized behaviour of systems. When there is imperfect and faulty mechanical operation in the real system, the simulation result is a prediction of energy consumption when the building performs as intended, rather than a reflection of the actual performance. When applying detailed simulation to analysis of component level energy use and estimation of energy savings from new strategies, simulated results can provide an adequate estimation or prediction from an engineering perspective. There might be a discrepancy between simulated results and real performance due to uncertainties.

CONCLUSION

Historical data based models are suitable for predicting whole-building energy use when there is sufficient but incomplete metered data. The inputs are easy to acquire and of little uncertainty. Model development requires affordable time, efforts and experience within the scope of an engineering project. Artificial neural network models provide relatively more accuracy for predictions on a short time scale than regression models. However, in this case study, regression models can achieve similar accuracy for predictions on a daily basis. There is little difference between monthly predictions of ANN models and regression models. Therefore, both regression and ANN models are suitable for predicting long-term whole building energy use, as far as the accuracy is concerned. Historical data based models are not able to give details on component level energy consumption. Detailed simulation models are able to give a relatively reliable analysis of component level energy use, if the models are well calibrated. When estimating energy savings of implemented retrofits, both regression and ANN models can be applied if there are sufficient data before and after the retrofits. However, regression and ANN models could not estimate energy savings potential of a new energy saving strategy, when there is no metered data after implementing the changes. Physical principle based models can be applied in this task. Table 2 is a

summary of recommended modeling methods for different engineering applications to existing buildings.

Eleventh International Conference on Building Simulation

 Table 2

 Recommended modeling methods for different

 engineering applications to existing buildings

Applicat	Recommended modeling method		
Whole-building	Short-term	ANN	
prediction	Long-term	Regression, ANN	
Component-level ene analysis	Detailed simulation		
Estimation of	Implemented retrofit	Regression, ANN	
energy savings	New strategy	Detailed simulation	

REFERENCES

- Cohen, D., Krarti. M. 1995. A neural network modeling approach applied to energy conservation retrofits. Paper presented at Proceedings of Fourth International Conference on Building Simulation.
- Dodier, R.H., Henze. G.P. 2004. Statistical analysis of neural networks as applied to building energy prediction. *Journal of Solar Energy Engineering* 126 : 592.
- Haves, P., Salsbury, T., Claridge, D., Liu, M. 2001. Use of whole building simulation in on-line performance assessment: Modeling and implementation issues.
- Kissock, J.K., Reddy, T.A., Claridge, D.E. 1998. Ambient-temperature regression analysis for estimating retrofit savings in commercial buildings. *Journal of Solar Energy Engineering* 120 : 168.
- Liu, M., Claridge, D.E. 1998. Use of calibrated HVAC system models to optimize system operation. *Journal of Solar Energy Engineering* 120 : 131.
- National Fire Protection Association (NFPA). Standard NFPA 45-2000. Standard on Fire Protection for Laboratories Using Chemicals, Section 6.3.3, p. 12
- Yan, B., Malkawi, A. M., Zhang, Y., Xia, J. 2009. Case study of energy diagnosis simulation of VAV AHU system controls. Proceedings of