Real-Time Forecast Model Analysis of Daily Average Building Load for a Thermal Storage System Control

Li Song, Ph.D., P.E.
School of Aerospace and Mechanical Engineering
University of Oklahoma

Ik-Seong Joo, Ph.D., P.E.

Subroto Guwana, Ph.D., P.E.
Ameresco Inc

Abstract

Thermal storage systems were originally designed to shift the on-peak cooling production to off-peak cooling production to reduce the on-peak demand. Based on the current electricity charging structure, the reduction of both on-peak and off-peak demands is becoming an exceedingly important issue. Reduction of both on-peak and off-peak demands can also extend the life span and defer or eliminate the replacement of power transformers due to potential shortage of building power capacity with anticipated equipment load increases. The next day daily average electricity demand is a critical set point to operate chillers and associated pumps at the appropriate time. For this paper, a mathematic analysis was conducted for annual daily average cooling of a building and three real-time building load forecasting models were developed. They are first-order autogressive model, random walk model and linear regression model. Finally, the comparison of results show the random walk model provides the best forecast.

Introduction

During past decades, many researchers have investigated the optimal control of the thermal storage systems to achieve the minimum operational cost and made significant progress (e.g., Tamblyn 1985, Braun 1990, Wei 2002, Massie and etc. 2004 and Liu and Henze 2007).

Among all the researchers, real-time forecasting of building load is critical for the thermal storage system optimization. Underestimation of the building load can cause unexpected chiller operation during on-peak hours and overestimation of the building load can overcharge the tank and generate extra heat loss through the storage tank. Henze [1997, 2004, 2005 and 2007] developed and tested a model-based predictive controller for optimal thermal storage systems control by adopting neural network theory into HVAC system control. In his research, he developed next 24 hours weather forecast models and used calibrated TYNSYS to simulate building performance. Inevitably, the internal heat gain, which has significant impacts on the building cooling load for commercial buildings, was considered as constant. In addition, it is not possible to adopt this method in the existing building automation system (BAS) due the computational requirements. Wei [2002] and Zhou [2005] developed practical optimization measures for thermal storage system control. In their analysis, the building load was projected simply by regression with the outside air temperature separated by weekdays and weekends. For the same outside air temperature, the highest building load was almost double of the lowest building load. The regression model versus outside air temperature cannot precisely describe the building load. In both studies, the occupancy schedule and building use changes were not taken into account.

Seem and Braun [1991] compared different algorithms for forecasting the building electrical loads in commercial buildings. An adaptive algorithm was proposed in their research. By defining the building electrical demand as a non-stationary time series, because of the fact that the electricity demand is dependent upon the day of the week and the time of the day, a combined model of CMAC (calculated using exponentially weighted moving average model) and autoregressive model (AR) is recommended, in which CMAC simulates the deterministic part and AR simulates the stochastic part. Linear interpolation based on minimum and maximum ambient temperature was used to incorporate the ambient temperature influence on the electrical demand. The accuracy of the combined CMAC and AR (3) model is verified to be acceptable. Because of the fact that the combined CMAC and AR (3) model doesn't need to store all the previous data, the computational and memory requirements for this method is relatively low.

The objective of the study in this paper is not to minimize the electricity demand during on-peak hours, but rather to reduce the electricity demand during both on-peak and off-peak hours and extend the lifetime or avoid the replacement of the existing transformer. The study should aim 1) to provide accurate forecasting of next 24 hours average daily electricity demand or next 24 hours total electricity consumption instead of daily profile and 2) to have the algorithm simple enough to be embedded in the existing BAS. By mathematic analysis, it was found that three models can be applicable for the forecasting. They are first-order autogressive model, random walk model and linear regression model. Finally, the comparison of results show the random walk model provides the best forecast. In a conclusion, through two steps of the model validation, even through all three models provide best fit for the historical data, the random walk model provides the best forecast by using new set of daily average data.

Average Building Load Forecast Model

Identifying the valid forecasting model is the most crucial in the application of predicting the next day's daily average cooling load for flattening the electrical demand in thermal storage operation. In this section, an annual average daily cooling load in a real facility is used to identify, develop and validate a forecast model for daily average cooling load. First of all, three different types of days must be defined for the analysis: unoccupied, occupied days when the previous day was occupied and occupied days when the previous day was unoccupied. This paper only presents the results of occupied days when the previous day was occupied.

Data Analysis

Actual measured annual cooling consumption data of a building located at Austin, Texas were collected for this model analysis. Figure 1 presents the daily average cooling consumption for the occupied days when the previous day was occupied. The data profile shows statistical behavior changes in time, and thus indicates that the daily average cooling is a non-stationary time series.

To prove the characteristics of the data, two concepts need to be introduced, autocorrelation function (ACF) and partial autocorrelation function (PACF) [Montgomery 2007]. To define ACF, autocovariance function at lag k needs to be defined first.

$$\gamma_k = Cov(y_t, y_{t+k}) = E[(y_t - \mu)(y_{t+k} - \mu)] \quad (1)$$

The collection of γ_k where k=0, 1, 2, … is called the autocovariance function

The autocorrelation coefficient at lag k is

$$\rho_k = \frac{E[(y_t-\mu)(y_{t+k}-\mu)]}{\sqrt{E[(y_t-\mu)^2 E(y_{t+k}-\mu)^2]}} = \frac{Cov(y_t,y_{t+k})}{Var(y_t)} = \gamma_k \gamma_0 \quad (2)$$

The collection of ρ_k where k=0, 1, 2, … is called the ACF.

The PACF is defined as the autocorrelation between y_t and y_{t-k} after adjusting for y_{t-1}, y_{t-2}, … , y_{t-k+1}.

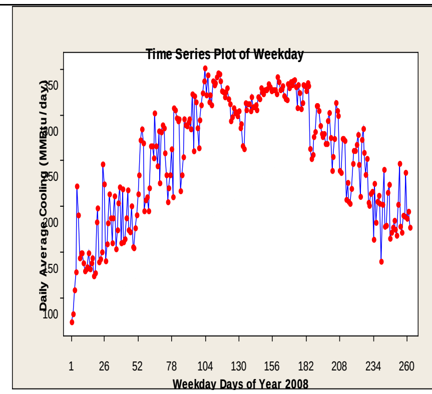

Figure 1: Time series plot of daily average cooling for weekdays

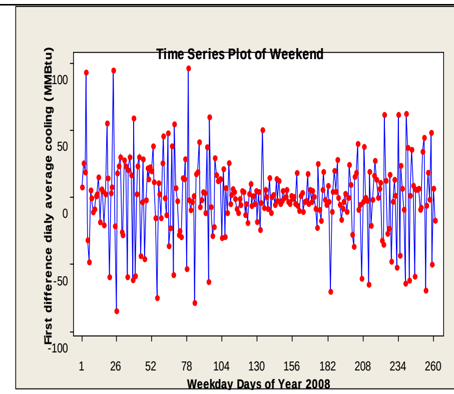

Figure 4: Time Series Plot of First Difference

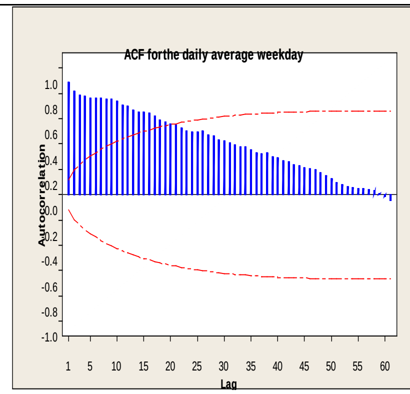

Figure 2: Sample ACF of daily average cooling for weekdays

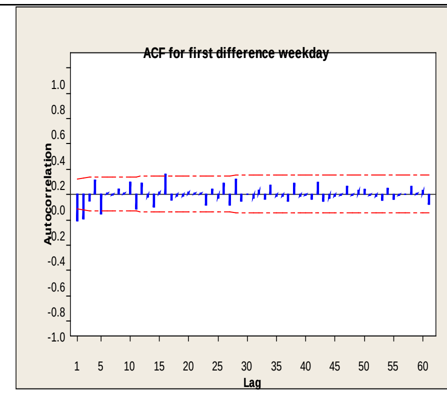

Figure 5: Sample ACF of the First Difference for weekdays

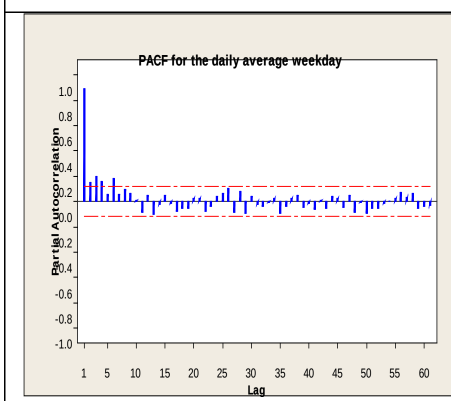

Figure 3: PACF of daily average cooling for weekdays

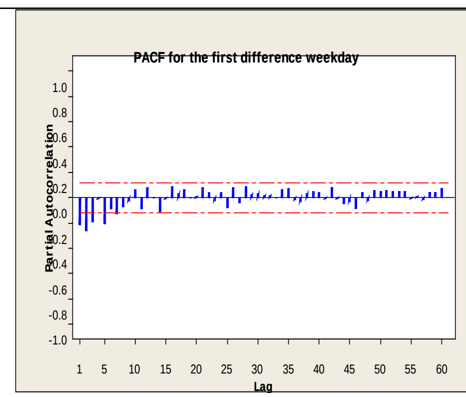

Figure 6: PACF of the First Difference for weekdays

Figure 2 and figure 3 are the sample ACF and PACF values for the daily average data. It can be seen in Figure 2 that sample ACF doesn't die out quickly, which leads to the conclusion that the daily average cooling is a non-stationary time series. In figure 3, the PACF chart, there is one significant sample PACF value, which suggests a first order autoregressive model (AR(1)) can fit the data [Montgomery 2007].

As a non-stationary time series, the daily cooling average data need to be tested if its first difference, that is , $w_t = y_t - y_{t-1} = (1-B)y_t$, or higher order differences, $w_t = (1-B)^d y_t$, produces a stationary time series. Figure 4 presents the time series plot of the first difference, which shows a very typical stationary behavior. The sample ACF and PACF of the first difference are plotted in Figure 5 and Figure 6 respectively. The ACF values are randomly positive and negative with values near zero, which represents a stationary behavior. By observing the ACF and PACF of the first difference, it is clear that the differencing the original data once eliminates the autocorrelation. Thus, the result suggests that the daily average data is less dependent on previous days' data. A random walk model ARIMA(0,1,0) can properly fit the data [Montgomery 2007] as well. ARIMA (p,d,q) is an autoregressive integrated moving average model with orders p, d and q. P represents Pth order of autoregressive, d represents dth difference, and q represents qth order of moving average.

Model Identification

Through previous data analysis, it is possible to use AR (1) and ARIMA (0,1,0) model to forecast the daily average cooling. The AR (1) model is given in equation (3).

$$\hat{y}_t = \delta + \varphi y_{t-1} + \varepsilon_t \qquad (3)$$

Where φ is a coefficient, δ is an constant and ε_t is noise at day t.

The ARIMA (0,1,0) model is given in equation (3).

$$\hat{y}_t - \hat{y}_{t-1} = \delta + \varepsilon_t \qquad (4)$$

The difference between the two models is by comparing equation (3) and (4) is the weighting factor (coefficient) for previous day in equation (3) that represents the influence on forecasting the next day's daily average cooling load.

In many buildings, the building thermal load is heavily dependent on weather. Integrated with physical knowledge about building cooling load, the combined value of constant (δ) and noise at day t (ε_t) is determined by future day outside air impacts on the cooling load. Therefore, equation (3) and (4) can be converted to following formats:

$$\hat{y}_t = C_1 y_{t-1} + C_2 \Delta y_{T_{oa},i} \qquad (5)$$

$$\hat{y}_t = y_{t-1} + \Delta y_{T_{oa},i} \qquad (6)$$

Where C_1 and C_2 are the weighting factors for previous day influence and temperature influence. $\Delta y_{T_{oa},i}$ represents the daily average cooling changes caused by outside air temperature.

By observing the time series chart of the first difference in figure 4, even though the mean of the first difference is around zero, it does show time changing variance. An online coefficient identification process of equation (5) is developed to determine the coefficients C_1 and C_2. The recursive identification method can provide another slow tracking on time varying system [L Ljung 1987]. Using this method, the next day's □*ton can be forecasted* from previous days' and previous weeks' data. Let *n* be number of data that we want to fit. For notational simplicity, let y_i be the *i*-th actual cooling load, and $\hat{y}_i$ be the predicted daily cooling load, *i*=1,…,*n*. Also let $y_{To,i}$ be cooling load changes caused by the outside air temperature. For optimal fit of the data, the sum of squares of the error must be minimal.

$$\underset{c_1,c_2}{\text{minimize}}: \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} y_i^2 - 2\sum_{i=1}^{n} y_i\hat{y}_i + \sum_{i=1}^{n}\hat{y}_i^2$$

$$\underset{c_1,c_2}{\text{minimize}}: \sum_{i=1}^{n} y_i^2 - 2c_1\sum_{i=1}^{n} y_i y_{i-1} - 2c_2\sum_{i=1}^{n} y_i y_{To,i} + c_1^2\sum_{i=1}^{n} y_{i-1}^2 + 2c_1c_2\sum_{i=1}^{n} y_{i-1}y_{To,i} + c_2^2\sum_{i=1}^{n} y_{To,i}^2$$

First-order derivatives for optimality dictate that the gradient at the stationary point is zero. Therefore:

$$\frac{\partial f}{\partial c_1} = -2\sum_{i=1}^{n} y_i y_{i-1} + 2c_1\sum_{i=1}^{n} y_{i-1}^2 + 2c_2\sum_{i=1}^{n} y_{i-1}y_{To,i} = 0 \quad \text{, and}$$

$$\frac{\partial f}{\partial c_2} = -2\sum_{i=1}^{n} y_i y_{To,i} + 2c_2\sum_{i=1}^{n} y_{To,i}^2 + 2c_1\sum_{i=1}^{n} y_{i-1}y_{To,i} = 0$$

Or in matrix notation: $\mathbf{AX} = \mathbf{B}$, where

$$\mathbf{A} = \begin{bmatrix} \sum_{i=1}^{n} y_{i-1}^2 & \sum_{i=1}^{n} y_{i-1}y_{To,i} \\ \sum_{i=1}^{n} y_{i-1}y_{To,i} & \sum_{i=1}^{n} y_{To,i}^2 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \text{and}$$

$$\mathbf{B} = \begin{bmatrix} \sum_{i=1}^{n} y_i y_{i-1} \\ \sum_{i=1}^{n} y_i y_{To,i} \end{bmatrix}.$$

The solution to this linear equation is $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$, where

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\text{adj}(\mathbf{A}) \text{ ,}$$

asured
ean
he

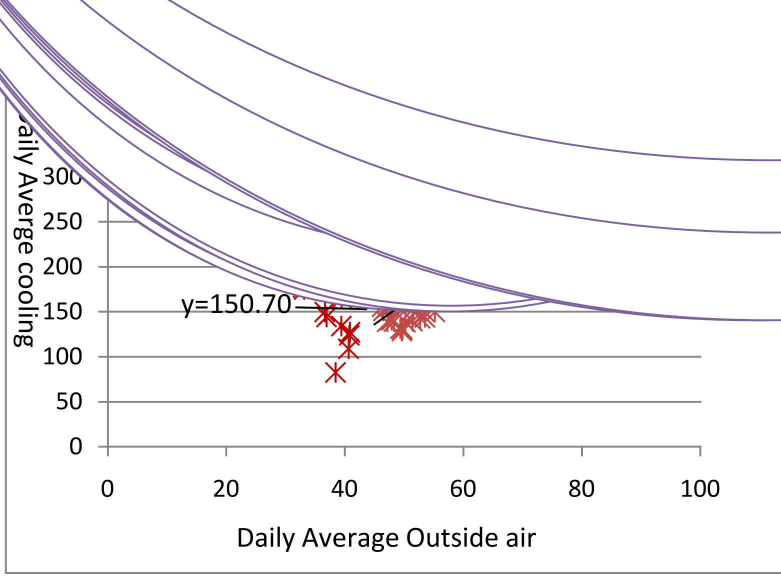

Figure 7: Daily Average Cooling versus Outside Air Temperature

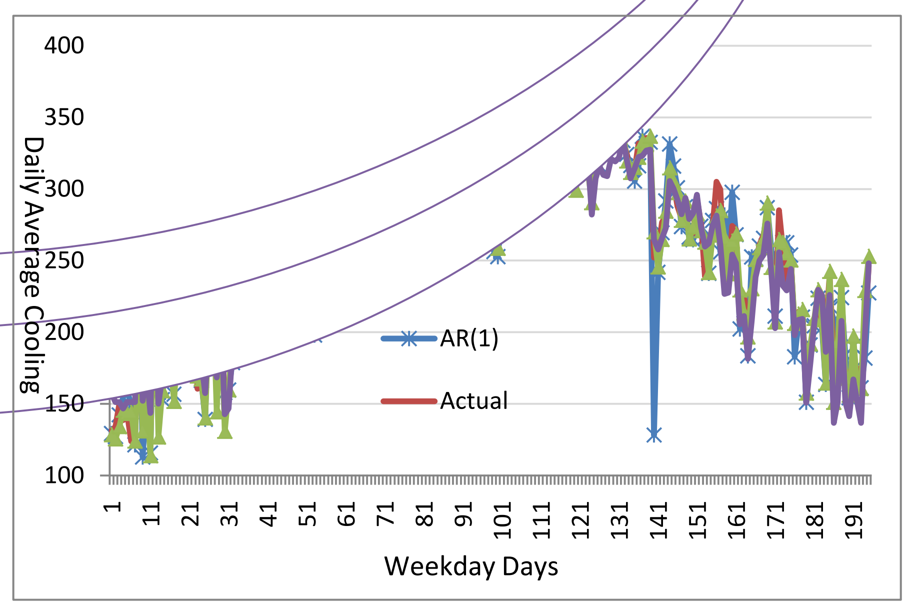

Figure 8: comparison of fitting results of three models with the actual values

st

on for
bution
sion
errors
usion,
hree models.

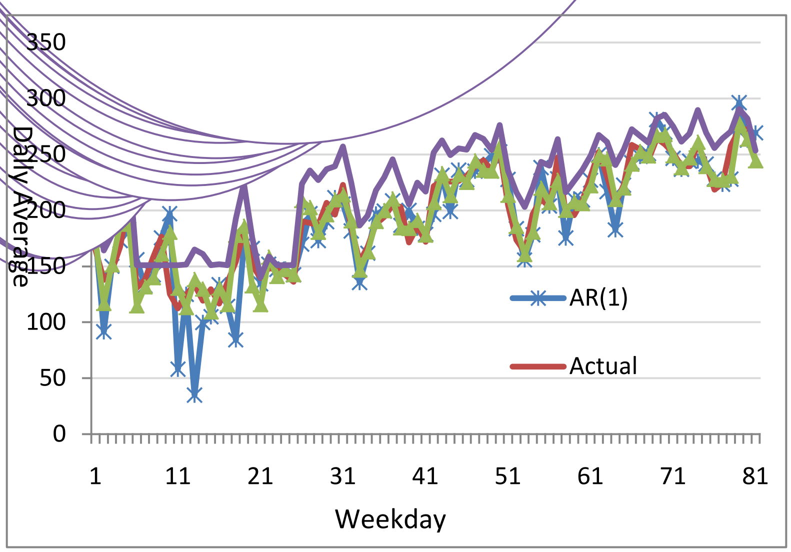

Figure 10: Comparison of forecasting results of three models and actual values

Table 2: Forecasting MAPE comparison of three models

Models	MAPE
Adaptive AR(1)	0.0900
ARIMA(0,1,0)	0.0593
Linear regression	0.1497

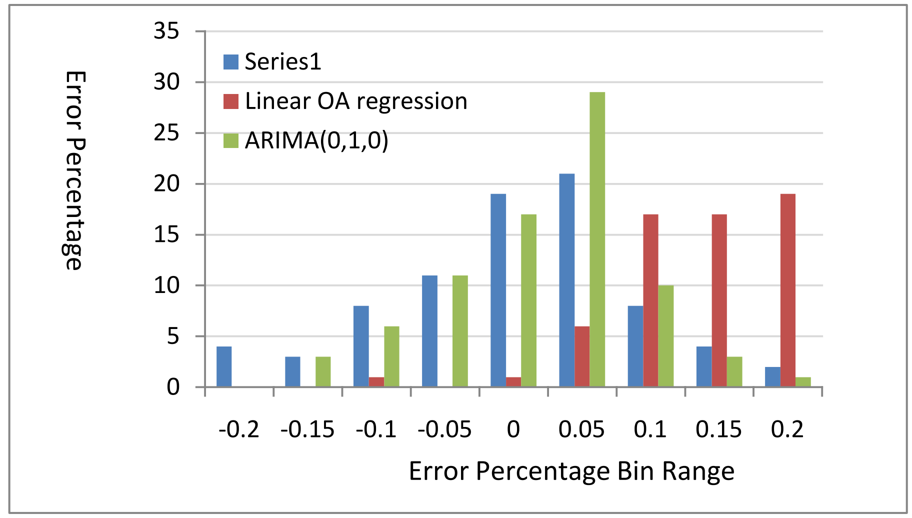

Figure 11: Comparison of forecasting error percentages of three models

Conclusions

To reduce both on-peak and off-peak demands and expand the life-span of the existing transformers with increasing building load due to building function changes or building expansion, a real time forecasting model is needed for projecting the next day daily average cooling for a thermal storage system control. Three models, first-order autoregressive model, random walk model and linear regression model, were developed through mathematic analysis using daily average cooling data for a building. With similar fitting accuracy of the three models, random walk model (ARIMA (0,1,0)) provides the best forecast results. The computation requirements for random walk model are relatively low and make it possible to be embedded in building automation system. An application of this algorithm will be presented in the next paper.

Reference

Tamblyn, R.T. 1985. Control concept for thermal storage, ASHRAE Transaction 91 (1B):5-11

Braun J.E. 1990. Reducing energy costs and peak electrical demand through optimal control of building thermal mass. ASHRAE Transaction 96(2):876-888

Wei, G., Liu, M.,Sakurai, Y., Claridge,D., Turner, W.D 2002. Practical optimization of full thermal storage systems operation, ASHRAE Transactions 108(2):360-368

Massie, D. D., Kreider, J. F., Curtiss, P. S., 2004 Verification of a Neural Network-Based Controller for Commercial Ice Storage Systems, ASHRAE Transactions vol. 110 (2):471-478

Liu, S., Henze, G. P. 2007 Evaluation of reinforcement learning for optimal control of building active and passive thermal storage inventory, Journal of Solar Energy Engineering vol129:215-225

Henze, G.P., M. Krati, and M.J. Brandemuehl. 1997a. A simulation environment for the analysis of ice storage controls. International Journal of HVAC&R Research 3(2):128-148

Henze, G.P., C. Felsmann, and G. Knabe. 2004 Evaluation of optimal control for active and passive building thermal storage. International Journal of Thermal Sciences 43(2):173-183

Henze G.P., Kalz D.E., Liu S., 2005 Experimental analysis of model-based predicative optimal control for active and passive building thermal storage inventory, HVAC&R research Volume 11(2): 189-214

Zhou, J., Wei, G., Turner W. D. 2005 Control Optimization for Chilled Water Thermal Storage System Under Complicated Time-of-Use Electricity Rate Schedule, ASHRAE Trans

J.E. Seem, J.E Braun, 1991 Adaptive methods for real-time forecasting of building electrical demand, ASHRAE Transaction vol.97(1):710-721

D.C. Montgomery, C. L. Jennings, M. Kulahci. 2007. Introduction to time series analysis and forecasting, ISBN 978-0-471-65397-4.

L Ljung, T. Söderström, 1987, Theory and application of recursive identification, ISBN 978-0-262-12095-1