Building Energy Use Prediction Using Time Series Analysis

ZHOU Ruijin
CISDI R&D Co., Ltd
Chongqing, China
ruijin.zhou@cisdi.com.cn

PAN Yiqun, HUANG Zhizhong, WANG Qiujian
Tongji University
Shanghai, China
yiqunpan@tongji.edu.cn

***Abstract*—This paper proposes a data-driven modeling method for building energy consumption prediction and applies it to two actual commercial buildings. Time series analysis is adopted as a main methodology to produce the data-driven model based on monthly actual energy consumption data. The models can be used to predict building future energy consumption, after being modified and verified.**

Keywords—Time Series; Combined Model; Energy Consumption Prediction

I. INTRODUCTION

In recent years, a large number of approaches for building energy prediction have been proposed and implemented in the process of building design, operation or retrofit from building's sub-system analysis[1-5]. Most data-driven models in heating, ventilation and air-conditioning (HVAC) field adopt either MLR (Multiple Linear Regression) analysis method or ANN (Artificial Neural Network) analysis method[6] - [12] [14]. Due to the strong inertia of building and HVAC systems, the building energy consumptions are related to the energy consumption during previous period with a random fluctuation . Therefore, time series analysis method can be used as an effective and accurate method to predict building energy consumption. However, This method was rarely studied.

This paper introduces an application of time series analysis method on the analysis and prediction of energy consumption of commercial buildings, so as to provide information for building energy conservation and management.

II. PREDICTION MODEL OF BUILDING ENERGY CONSUMPTION BASED ON TIME SERIES ANALYSIS

Based on the statistical research on the time series data of each variable, this paper proposes an energy consumption prediction method applied to commercial buildings, by adopting time series analysis method combined with regression method. By using the historical monthly energy consumption data of the previous three years, the annual building energy consumption of the fourth year can be predicted with proposed method. Building energy consumption prediction model can be established in four steps:

A. Original Sequence and Seasonal Adjustment.

First, set up a building energy consumption sequence $\{X_t\}$. Because some objective laws of the sequence can be covered by the seasonal fluctuation of monthly data, perform seasonal adjustment to the original sequence is necessary by adopting CensusX12 method addition model. By employing the software of Eviews, the original sequence $\{X_t\}$ is decomposed into three parts, i.e. Trend Sequence (TC_t), Seasonal Factor (S_t) and Irregular Factor(I_t), which is a complicated iterative procedure:

$$X_t = TC_t + S_t + I_t \quad (1)$$

Then Trend Sequence (TC_t) can be obtained by rearranging Equation (1):

$$TC_t = X_t - S_t - I_t \quad (2)$$

Denominate the sum of Seasonal Factor (S_t) and Irregular Factor(I_t) as SI_t:

$$SI_t = S_t + I_t \quad (3)$$

Combine Equation (1) and Equation (3) to get:

$$X_t = TC_t + SI_t \quad (4)$$

B. Pre-process, Model establishment and Trend Sequence(X_t) Prediction

Before establishing model, the Trend Sequence (TC_t) should be conducted with stationary test. Then various methods are adopted to build a time-series model according to the test results, i.e. establishing time-series model directly if they are stationary; adopting combined model method if they are non-stationary. The Pure Randomness Test is also required before modeling by adopting Q Statistic Test method. The time-series model can be established only for the sequence which passes two mentioned tests after the process of model identification, order determination, parameter estimation and adaptability test.

The Non-stationary TC_t Sequence consists of a trend part and a random part. The combined model is established in three steps：

1) Fit the trend part into a function according to sequence characteristics, see Equation (5), and make Residual Sequence Y_t stationary;

2) Establish the optimal time-series model of Y_t Sequence, which passes Stationary Test and Pure Randomness Test, and predict the future value of Y_t Sequence with the result model [15];

3) Combine the established models of trend part and Residual Sequence Y_t to obtain the final combined model of TC_t Sequence and predict the future value, see Equation (6)：

$$TC_t = f(t) + Y_t \tag{5}$$

$$\begin{aligned} TC_t &= f(t) + Y_t \\ &= f(t) + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_n Y_{t-n} \\ &\quad + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_m a_{t-m} \end{aligned} \tag{6}$$

C. Physical Regression and Prediction of SI_t Sequence

By analyzing and predicting building energy consumption (i.e. electricity, oil and gas), Time-series model can be simplified and the energy management and energy conservation measures can be implemented more conveniently. Four impact factors are taken into account, i.e. Monthly Accumulated Temperature Value ($TEMP_t$), Monthly Average Relative Humidity (RH_t), Monthly Workdays (WD_t), and Monthly Non-workdays (NWD_t). The monthly rental rate of commercial building involves commercial secrets, the precise value of which is difficult to be acquired. Therefore, this paper just studied the energy consumption prediction method for buildings with basically constant occupancy rates. The method proposed in this paper has limitations for the buildings with variable occupancy rates [16].

It is more practical and cost-effective to study the seasonal factors and irregular factors by using fitting method and expressing SI_t Sequence as a regression function, see Equation (7). Then it can be used to predict electricity consumption [16]:

$$\begin{aligned} SI_t &= c + a_1 TEMP_t^2 + a_2 RH_t^2 + a_3 WD_t^2 + a_4 NWD_t^2 \\ &\quad + a_5 TEMP_t \cdot WD_t + a_6 TEMP_t \cdot NWD_t + a_7 RH_t \cdot WD_t \\ &\quad + a_8 RH_t \cdot NWD_t + a_9 TEMP_t + a_{10} RH_t + a_{11} WD_t + a_{12} NWD_t \end{aligned} \tag{7}$$

Since the latent heat load accounts for a very small proportion of the entire load during winter, RH_t will not be included as an input in the prediction model for building gas or oil consumption as below

$$\begin{aligned} SI_t &= c + a_1 TEMP_t^2 + a_2 WD_t^2 + a_3 NWD_t^2 + a_4 TEMP_t \cdot WD_t \\ &\quad + a_5 TEMP_t \cdot NWD_t + a_6 TEMP_t + a_7 WD_t + a_8 NWD_t \end{aligned} \tag{8}$$

Software Eviews was selected to fit SI_t sequence to establish a physical regression model, and predict the future value of SI_t for 2010 by inputting the predictive values of these four independent variables.

The accumulated temperature of a certain month of each year fluctuates little, which is related to the values of that month of previous years. For the prediction of $TEMP_t$, different weights are taken for the terms of previous three years as below:

$$TEMP_{2010,i} = b_1 TEMP_{2009,i} + b_2 TEMP_{2008,i} + b_3 TEMP_{2007,i} \quad (i=1，2，…，12) \tag{9}$$

b_1 , b_2 , b_3 are the weight factors for the terms of 2009, 2008, 2007 respectively, and $b_1 + b_2 + b_3 = 1$. The closer of the year, the more influence on the values of predictive accumulated temperature. After comparing the accuracy of prediction under 20 groups with different weight factors, the group of $b_1 = 0.5$, $b_2 = 0.35$, $b_3 = 0.15$ is chosen. So $TEMP_t$ can be calculated based on the following equation:

$$TEMP_{2010,i} = 0.5 TEMP_{2009,i} + 0.35 TEMP_{2008,i} + 0.15 TEMP_{2007,i} \quad (i=1，2，…，12) \tag{10}$$

Meanwhile, the prediction of Monthly Average Relative Humidity can be conducted in the same way as the accumulative temperature.

$$RH_{2010,i} = 0.5 RH_{2009,i} + 0.35 RH_{2008,i} + 0.15 RH_{2007,i} \quad (i=1，2，…，12) \tag{11}$$

The future value of WD_t and NWD_t can be calculated according to Chinese calendar. Thus, the SI_t sequence can be predicted from the equation below:

$$\begin{aligned} SI_{2010,i} &= f(Temp_{2010,i}, RH_{2010,i}, WD_{2010,i}, NWD_{2010,i}) \\ &= f(0.5 Temp_{2009,i} + 0.35 Temp_{2008,i} \\ &\quad + 0.15 Temp_{2007,i}, 0.5 RH_{2009,i} + 0.35 RH_{2008,i} \\ &\quad + 0.15 RH_{2007,i}, WD_{2010,i}, NWD_{2010,i}) \end{aligned} \tag{12}$$

D. Modification for Heating Energy Consumption

In the last step, the predicted values of monthly gas consumption and monthly oil consumption should be adjusted, because the prediction for the months without heating demand may be unequal to zero or even negative. The adjustment is done by two steps: 1. Adjusting the negative values as zero ; 2.Adjusting the values of the months as zero, when neither gas nor oil is consumed (e.g. from April to October) according to the actual historical energy consumption of that project in the past years. The final predicted results can be acquired after adjustment.

III. Case Analysis

This study applies the building energy prediction theory and method on the future load analysis for two buildings. Both buildings have nearly constant occupancy rate, but different heating & cooling energy source types, as shown in TABLE I

TABLE I. ENERGY SOURCE TYPE AND OCCUPANCY RATE OF TWO CASES

Case	Building Energy Sources	Occupancy Rates
1	chiller & gas boiler	nearly constant
2	chiller & oil boiler	nearly constant

A. *Case 1—Office Building A in Shanghai*

The total floor area of the building is 77000 m^2, about 80% of which is air conditioning area

- Building Electricity Consumption Prediction

Step 1: collect the monthly electricity consumption data from 2007 to 2009 (36 data points in total), as shown in Figure 1. These data present obvious seasonal characteristics. Then, the seasonal adjustment is done using X12 addictive model.

Step 2: implement Stationary Test on $Electricity\_TC_t$ Sequence. The test results show that P value is much greater than 0.05, which means there is unit root in the sequence and thus $Electricity\_TC_t$ Sequence is non-stationary. So the combined model is applied:

$$Electricity\_TC_t = f(t) + Electricity\_Y_t \tag{13}$$

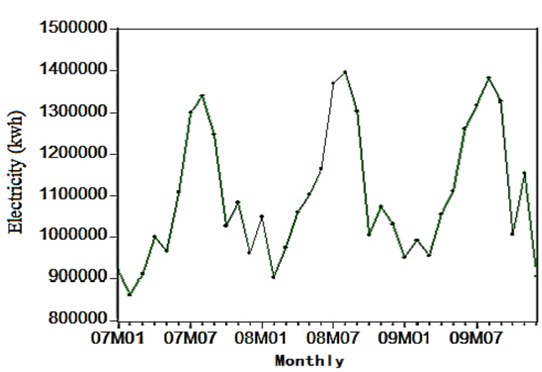

Figure 1. Monthly electricity consumption sequence of office building A in Shanghai

Firstly, employ SPSS to fit the $Electricity\_TC_t$ sequence and eventually choose cubic function as the optimal function type for the trend part after trial and error of function types. Secondly, implement Stationary Test and Pure Randomness Test on Residual Sequence $Electricity\_Y_t$. The statistical result shows the tests are passed and the next step can be performed. Thirdly, ARMA model is established for Residual Sequence $Electricity\_Y_t$. 6 models were inputted into the fitting trials and the model of ARMA(2,3) was selected as the optimal method based on the order determination method.

After combining the trend part model $f(t)$ and the residual model $Electricity\_Y_t$, the final combined predictive $Electricity\_TC_t$ sequence is obtained as below:

$$\begin{aligned} Electricity\_TC_t &= f(t) + Electricity\_Y_t \\ &= 3.095t^3 - 373.572t^2 + 12170.963t \\ &\quad +999718.5 + 1.654Electricity\_Y_{t-1} \\ &\quad -0.835Electricity\_Y_{t-2} + a_t \\ &\quad +1.223a_{t-1} + 0.549a_{t-2} - 0.293a_{t-3} \end{aligned} \tag{14}$$

Step 3: build the physical regression model for $Electricity\_SI_t$ and predict the future values. Four groups of dependent variable data points are used to fit $Electricity\_SI_t$ sequence according to Equation (8). The value of R^2 is 0.902 and P value is 0.0000. This means that $Electricity\_SI_t$ sequence can be related to above four variables. The $Electricity\_SI_t$ sequence can be calculated as Equation (15):

$$\begin{aligned} Electricity\_SI_t &= 9974635.9 + 1.1TEMP_t^2 + 184.4RH_t^2 \\ &\quad -5003.7WD_t^2 + 4525.2NWD_t^2 \\ &\quad -107.9TEMP_t \cdot WD_t - 96TEMP_t \cdot NWD_t \\ &\quad +4679.5RH_t \cdot WD_t + 5673.1RH_t \cdot NWD_t \\ &\quad +2501.7TEMP_t - 177168.8RH_t \\ &\quad -121996.4WD_t - 502899.6NWD_t \end{aligned} \tag{15}$$

Then, the monthly accumulated temperature and the average relative humidity of 2010 in Shanghai are predicted from Equation (10) and Equation (11). The monthly workdays and the non-workdays of Office Building A in 2010 are calculated according to Chinese calendar. All these values are introduced into Equation (15) to calculate $Electricity\_SI_t$ sequence.

Step 4: predict the monthly electricity consumption of Office Building A. $Electricity\_TC_t$ of 2010 can be calculated from Equation (14), then the total electricity consumption can be calculated by adding $Electricity\_SI_t$:

$$Electricity_t = Electricity\_TC_t + Electricity\_SI_t \tag{16}$$

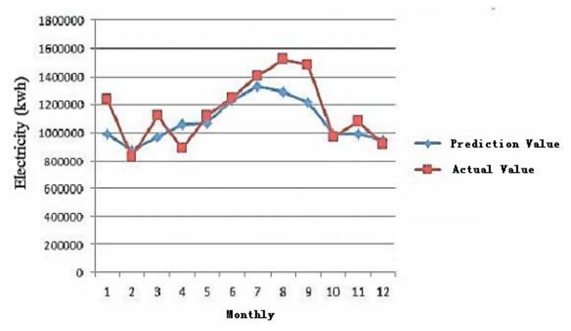

Figure 2. Prediction vs. actual value of electricity consumption of office building A

The predicted results are shown in Figure 2. Figure 2 shows that the predicted values generally follow the actual values. The annual electricity consumption error is -6.326%.

- Building Gas Consumption Prediction

The gas consumption prediction model is obtained by adopting the same method as electricity consumption prediction model, except the second step of model establishment. The stationary test results show that Gas_TC_t Sequence is stable. Therefore, the time series model is established for prediction directly.

By comparing the gas consumption data of the previous three years, the negative values should be modified to 0. The modified values of gas consumption prediction for Office Building A are shown in Figure 3. Figure 3 shows that the predictive values follow the actual values well and the annual gas consumption error is -1.753%.

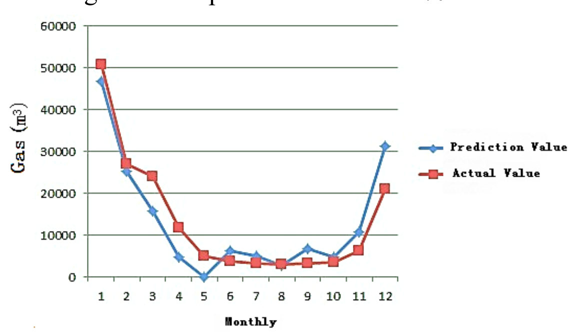

Figure 3. Prediction vs. actual value of gas consumption of office building A

B. Case 2—Commercial Building B in Shanghai

The total floor area of the building is 67000m^2, 91% of which is air conditioning area. The monthly electricity consumption data from 2007 to 2009 (36 data points in total) are collected to predict the energy consumption of 2010.

- Building Electricity Consumption Prediction

The electricity consumption prediction model is established with the same method as for Office Building A. As shown in Figure 4, a good fitting is indicated in the predicted results. The annual electricity consumption error is 1.0221%.

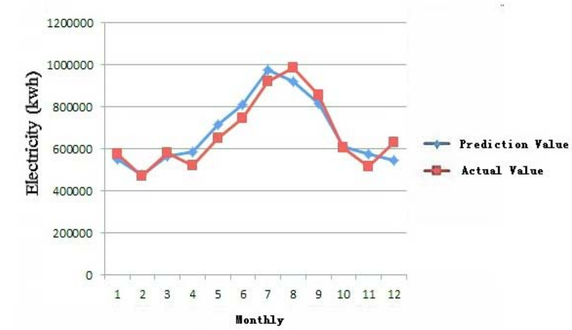

Figure 4. Prediction vs. actual value of electricity consumption of commercial building B

- Building Oil Consumption Prediction

The oil consumption prediction model is established with the same method as gas prediction of Office Building A. Figure 10 shows the predicted results. The data of oil consumption for non-heating seasons (May to October) should be modified to 0, shown in Figure 5. The modified predicted values follow the actual values well, and the annual oil consumption error is 5.885%.

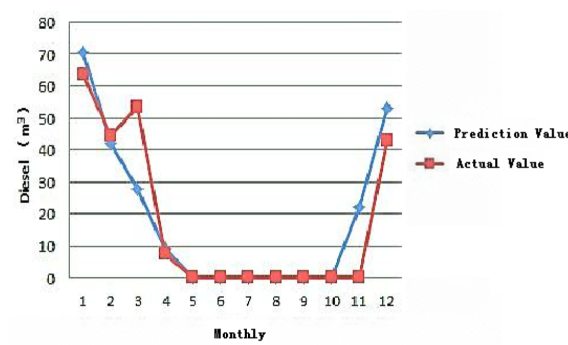

Figure 5. Prediction vs. actual value of oil consumption of commercial building B

IV. CONCLUSION

Based on the historical data and the physical principles, this study adopts a time series method on commercial building energy consumption analysis and predictions, presents a modeling method for building energy consumption prediction, and applies the method to different commercial buildings. After establishing, verifying, analyzing and evaluating the prediction models for building energy consumption, the following conclusions can be drawn:

- The time series analysis is an effective and accurate method for building energy consumption prediction. Due to the inertia of building system, the energy consumptions of current period are related to the previous period with some certain random changes
- According to the physical principles, the physical regression model is more practical. In this study, four main impact factors are introduced into the building energy consumption models, i.e. Monthly Accumulated Temperature Value, Monthly Average Relative Humidity, Monthly Workdays, and Monthly Non-workdays .
- This paper verified the feasibility of the building energy consumption prediction method based on time series and applied the method to two real cases. The results demonstrate the accuracy of the method.
- The building with variable occupancy rates is beyond the scope of this research, the prediction method presented in this paper can only be applied to commercial buildings with nearly constant occupancy rates. Meanwhile, the method is used to predict the total annual energy consumption of the building, not the energy consumption breakdown.

REFERENCES

[1] White JA, Reichmuth R. Simplified method for predicting building energy consumption using average monthly temperatures. In: Proceedings of the 31st intersociety energy conversion engineering conference, vol. 3. 1996.p. 1834–9.

[2] Yao R, Steemers K. A method of Equationting energy load profile for domestic buildings in the UK. Energy and Buildings 2005;37(6):663–71.

[3] Yik FWH, Burnett J, Prescott I. Predicting air-conditioning energy consumption of a group of buildings using different heat rejection methods. Energy andBuildings 2001;33(2):151–66

[4] Lei F, Hu P. A baseline model for office building energy consumption in hot summer and cold winter region. In: Proceedings of international conference on management and service science. 2009. p. 1–4

[5] Lam JC, Wan KKW, Wong SL, Lam TNT. Principal component analysis and long-term building energy simulation correlation. Energy Conversion and Management 2010;51(1):135–9.

[6] Jiao WenLing, Jin Jiabin, Lian Leming, Yan Mingqing. Application of time series analysis to load forecast of natural gas in town. Journal of Harbin University of Civil Engineering and Architecture.2001, 34(4)

[7] Zhou Enze, Fang Xiumu. The study on heat load forecast of the heating boiler house. Journal of Harbin University of Civil Engineering and Architecture.2000,33(3).

[8] S.Sp. Pappas, L. Ekonomou, P. Karampelas, D.C. Karamousantas, S.K. Katsikas, G.E. Chatzarakis, P.D. Skafidas. Electricity demand load forecasting of the Hellenic power system using an ARMA model. Electric Power Systems Research 80 (2010) 256-264

[9] Jin Yang, Hugues Rivard, Radu Zmeureanu, On-line building energy prediction using adaptive artificial neural networks. Energy and Buildings 37 (2005) 1250–1259

[10] T. Olofsson*, S. Andersson. Long-term energy demand predictions based on short-term measured data. Energy and Buildings 33 (2001) 85-91

[11] Cristian Ghiaus. Experimental estimation of building energy performance by robust regression. Energy and Buildings 38 (2006) 582–587

[12] Abdullatif E. Ben-Nakhi, Mohamed A. Mahmoud. Cooling load prediction for buildings using general regression neural networks. Energy Conversion and Management 45 (2004) 2127–2141

[13] Yiqun Pan, Zhizhong Huang, Xiaowei Zheng. Data Processing and Data Mining On Energy Consumption Database of Commercial Buildings in Shanghai. 2009

[14] Bin Yan, Ali M. Malkawi, Yun Kyu Yi. Case Study of applying different energy use modeling methods to an existing building. 12th Conference of International Building Performance Simulation Association, Sydney, 14-16 November

[15] Kissock, J.K., Reddy, T.A., Claridge, D.E. Ambient-temperature regression analysis for estimating retrofit savings in commercial buildings. Journal of Solar Energy Engineering (1998) 120-168.

[16] Joseph C. Lam, Kevin K.W. Wan, S.L. Wong, Tony N.T. Lam. Principal component analysis and long-term building energy simulation correlation. Energy Conversion and Management 51 (2010) 135–139