An improved ARX model for hourly cooling load prediction of office buildings in different climates

Chengliang Fan[1], Yundan Liao[1,2], Yunfei Ding[1,2,a]

[1]School of Civil Engineering, Guangzhou University, Guangzhou 510006, China
[2]Guangdong Provincial Key Laboratory of Building Energy Conservation and Application Technology, Guangzhou 510006, China

Abstract. An attempt was made to develop an improved autoregressive with exogenous (ARX) model for office buildings cooling load prediction in five major climates of China. The cooling load prediction methods can be arranged into three categories: regression analysis, energy simulation, and artificial intelligence. Among them, the regression analysis methods using regression models are much simple and practical for real applications. However, traditional regression models are often helpless to manage multi-parameter dynamic changes, making it not accurate as the other two categories. Many of the existing cooling load prediction studies use piecewise linearization to manage nonlinearity. To improve the prediction accuracy of regression analysis methods, higher order and interaction terms are included in improved ARX based on traditional ARX model. The improved ARX model consists of eight variables, with eleven coefficients accessed at a time. For applications and evaluations, an office building in major cities within each climatic zone was selected as a representation. These cities were Harbin, Beijing, Nanjing, Kunming and Guangzhou respectively. The coefficient of determination R^2 is greater than 0.9 in five cities. The prediction results show that the improved ARX model can adapt to different climatic conditions, including those nonlinearity cases.

1 Introduction

Building energy is a widely studied topic owing to its significant impact on global energy demand. In office buildings that have large energy intensity, the heating, ventilating, and air-conditioning (HVAC) system accounts for a proportion of 68% [1]. The reduction of building energy has been one of the primary topics for many developing and developed countries. Advanced control and optimal operation offer a potential to minimize HVAC system energy consumptions and significantly save operation costs by optimizing energy demand and supply in application. Building cooling load prediction has been regarded as one critical important technology in optimal operation to reduce HVAC system energy consumptions [2][3].

A substantial number of methods have been developed for building cooling load prediction. The methodologies on cooling load prediction can be broadly arranged in three categories: regression analysis, energy simulation, and artificial intelligence (AI) analysis. Energy simulation softwares, which includes DOE-2, ESP-r, TRNSYS, EnergyPlus, are capable of simulating building cooling load by using details of building including construction materials, windows performance, occupancy, etc. Meanwhile, meteorological data are needed, and its completeness and accuracy affect the simulation results directly [4]. In practice, it is hard to build up a building model with identical information to

reality in the software, because not all of the input variables, parameters, and setting that relate to the cooling load of HVAC systems can be determined correctly and measured accurately [5]. Meanwhile, simulation softwares cost considerable manpower and time to build and operate the simulation models. It is not widely used in practice to predict the cooling load of HVAC systems [4].

One of most popular AI methodologies is artificial neural network (ANN) [6], which has demonstrated better cooling load prediction accuracy than conventional regression methods [7]. A major disadvantage of ANN is that a large number of historical cooling load and weather data are needed to finish training. Moreover, the training process needs a large amount of time. The theory of AI is too complicated for engineers. ANN is more suitable for theoretical research but may not practical for real applications given to the complicated models and uncertainties.

Regression methods aim to predict building cooling load by finding out appropriate coefficients, which associated with the most influential inputs or historical cooling load. Compared with the AI and energy simulation, the easy operation of the regression analysis is one of the most important advantages in practice. Regression methods tend to be simple and easily applicable to any type of building [8]. At present, the

---

[a] Corresponding author: dingyf@126.com

main forms of regression analysis applied are multiple linear regressive (MLR) models [9], auto regressive (AR) models , and autoregressive with exogenous inputs (ARX) models [10] have been used as building cooling load prediction models.

Compared with other categories, although current regression analysis in literature is considered to be less accuracy, it is practical and simple to implement for real applications in building cooling load prediction. However, due to traditional regression analyses is often helpless to manage multi-parameter dynamic changes. Therefore, this paper developed an improved ARX model to improve the prediction accuracy of traditional ARX model. The higher order term of temperature and interaction term among exogenous variables in improved ARX model to model the nonlinearity. These terms can make improved ARX model adapt nonlinear condition and provide the capability to reduce the influence of outliers. The work involved four main aspects in this study:

(i) Generation of cooling load database through a series of EnergyPlus simulation runs for office buildings in major climate zones in China.

(ii) Develop an improved ARX model as functions of the key design variables using regression technique.

(iii) Application of the improved ARX model for the selected buildings in five climates.

2 Descriptions of major climates and building designs

China has a vast territory and a complex terrain. These diversity and complexity lead to different climate characteristics in many different regions. Areas with similar climate characteristics are divided into the same climate zone. The zoning standard is based on the average temperature of the coldest and hottest months of the year. Days when the daily average temperature is below 5℃ or above 25℃ are considered as a supplementary indices for determining the zones [11]. In terms of the thermal design of buildings, there are five major climates, namely severe cold, cold, hot summer and cold winter, temperate, and hot summer and warm winter in China [12]. This simple climate classification mainly involves conductive heat gain or heat loss and corresponding thermal insulation issues. A major city within each of the five climatic zones was selected for this study. These were Harbin (severe cold), Beijing (cold), Nanjing (hot summer and cold winter), Kunming (temperate) and Guangzhou (hot summer and warm winter) respectively. The distribution of the five thermal zones is shown in the Fig. 1.
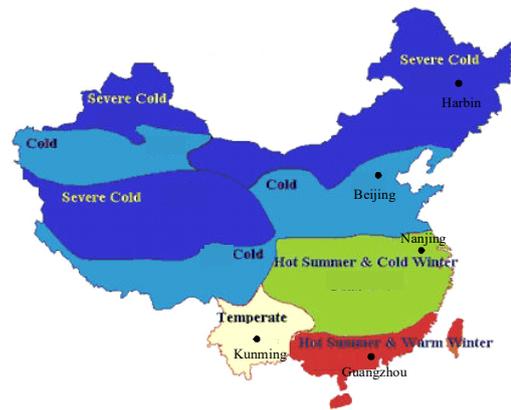

Figure 4. The different climates in China.

EnergyPlus is used to simulate the building hourly cooling load based on CSWD weather data. CSWD weather data consisting 8760 hourly records of dry bulb temperature, relative humility, solar radiation, wind speed and other meteorological parameters for each city was developed for the simulation exercise. EnergyPlus has been researched and proven to have extremely high calculation accuracy and reliability [13]. A base-case office building was designed to serve as a baseline reference for comparative cooling load prediction studies. The base-case was 70.8m×18.9m, 11-storey building, 3m floor-to-floor height, 0.4 window-to-wall ratio. The total floor area is 14719m² (12564m² air-conditioned, 2155m2 non-air-conditioned). Light power and equipment power are 20 W/m², 13 W/m² respectively. Occupancy density is 10 m²/person. The specifications of the interior design conditions are shown in Table 1. Table 2 shows the office building envelope basic-set parameters values for five cities. In order to compare the accuracy of cooling load prediction, it is assumed that the internal loads, the interior design conditions and the basic HVAC system have corresponding design data, which are derived from the local design specifications of China.

Annual air-conditioning operating cycle is from May to October. The operating time for air-condition system is from 8:00 to 19:00, 12-h of each weekday. The cooling is provided by cold water chiller. The summer indoor design temperature is 26℃.

Table 1. Description of the interior design conditions.

Light average usage	Until[a] (fraction[b]): 7 (0.05), 8 (0.3), 17 (0.9), 18 (0.3), 20 (0.2), 24 (0.05)
Equipment average usage	Until[a] (fraction[b]): 8 (0.4), 12 (0.9), 14 (0.8), 17 (0.9), 18 (0.5), 24 (0.4)
Occupancy schedule	Until[a] (fraction[b]): 7 (0.05), 8 (0.3), 12 (0.9), 14 (0.5), 17 (0.9), 18 (0.3), 20 (0.2), 24(0.05)

[a] The hour of the day until specified fraction is considered.
[b] The fraction of total value of the variable that is considered in

calculation for that specific period of time.

$[Load_1 \ Load_2 \ … \ Load_i]^T$. The weight coefficient is shown in Eq. (2).

$$W = (A^T \cdot A)^{-1} \cdot A^T \cdot L \qquad (2)$$

4 Results

The time span of cooling load is in July. The training cooling load data is compared with EnergyPlus simulated data of each city. The proposed matrix 'W' of prediction scheme is implemented by Eq. (2). The accuracy of proposed model in four cities is evaluated using the following statistical criteria: root mean squared error (*RMSE*), expected error percentage (*EEP*), coefficient of variation (*CV*) and mean absolute error (*MAE*) [3].

An example of the weight coefficients and relevant statistics of 120-h (five consecutive weekdays.) is shown in Table 3. It can be seen that the coefficient of determination R^2 varies from 0.906 in Beijing to 0.995 in Guangzhou. The R^2 values was estimated to quantify the curve-fitting accuracy of prediction cooling load within a time span, from 0 to 1, and higher R^2 values indicate a better fit. It seems that warmer climates in the south tend to have strong correlation between the building hourly cooling load and the dynamic variables with smaller *RMSE* values in training.

A low training error can not guarantee a low prediction error in practice. To verify that the weight coefficient values obtained by training were still used to valid in different climates. Related air-conditioning operation settings were described in Section 2. For clear presentation, the improved ARX model was validated using time span from August 27 to August 31. The results are shown in Figure 2-6. As can be seen, all five cities prediction cooling load show both underestimation and overestimation. The percentages of relative error in different cities are described. The deviation percentages are largely within ±10% according to the relative errors for the same time span in different cities. Table 4 shows the comparison of statistical evaluation in five cities. The R^2 values are from 0.920 in Beijing to 0.942 in Kunming, very closed. It is envisaged that the improved model has good adaption to predict building cooling load in different cities. The difference of *EEP* and *CV* evaluation indices values is not very obvious. *RMSE* and *MAE* values are different degrees of changes in these cities, such as *RMSE* and *MAE* values are 76.3kW and 28.5kW in Guangzhou, 43.3kW and 16.7kW in Harbin respectively. The cause of this might be external weather variables (such as Temp, Rad and RH) and internal gains (such as human activities, Occ, LP and EP) are different levels with different weather condition

Table 2. Building envelope basic-set parameters for 5 cities.

City	Wall	Window	Roof	Shading coefficient
	U-value (W/m²K)			
Harbin	0.44	2.50	0.35	0.64
Beijing	0.60	2.60	0.55	0.70
Kunming	1.47	3.50	0.89	0.55
Nanjing	1.00	3.00	0.70	0.60
Guangzhou	1.56	4.00	0.54	0.40

3 Description of improved ARX model

Two group variables [14] that associated with building dynamic cooling load can be categorized as follows:

(1) External weather: dry bulb ambient temperature (Temp), relative humidity (RH), direct radiation (Rad).

(2) Internal gains: people occupancy (Occ), lighting power (LP), equipment power (EP), ventilation rate (VR).

Based on traditional ARX model, to improve prediction accuracy of hourly cooling load, higher order and interaction terms are included in improved ARX to model nonlinearity and inter-correlation between input variables in this study. The improved ARX model is shown in Eq. (1).

$$\begin{aligned} Load_t &= w_1 \cdot Load_{act,t-1} + w_2 \cdot Temp_t + w_3 \cdot Temp_t^2 \\ &+ w_4 \cdot RH_t + w_5 \cdot Temp_t \cdot RH_t + w_6 \cdot Rad_t \\ &+ w_7 \cdot Occ_t + w_8 \cdot VR_t + w_9 \cdot LP_t + w_{10} \cdot EP_t \\ &+ w_{11} \end{aligned}$$

$$(1)$$

where 'w' is weight coefficient for input variables, 't' is predicted time. The initial weight coefficient matrix 'W' can be found as $[w_1 \ w_2 \ …w_{11}]^T$. 'W' could be modified and updated when predicted cooling load data has relatively big error with measured cooling load data, and the allowable error can be determined by data set. 'A' is a matrix for all input variables and the historical cooling load. 'A' can be written as $[x_1, x_2,…,x_{10}, 1]$, $x_1,x_2,…,x_{10}$ are 10 vectors of input parameters. The subscript i=11,12,13,…,n, and matrix $[A^T \cdot A]$ is invertible. 'L' is a matrix of historical cooling load. 'L' can be written as

Table 3. An example of the weight coefficients and relevant statistics for the 5 cities.

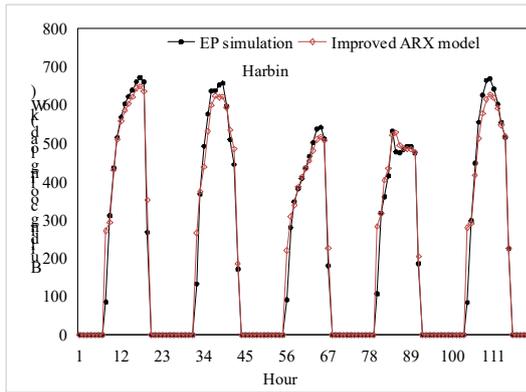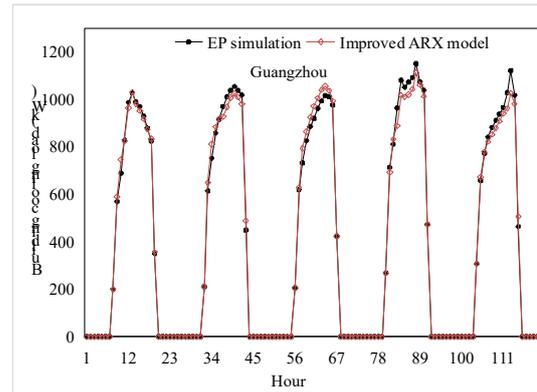| City	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}	R^2	*RMSE*
Harbin	0.54	-111.7	1.94	-10.5	0.51	0.09	-9.5	-2.51	17.4	20.3	1261.4	0.912	62.6
Beijing	0.60	-8.0	0.51	4.70	-0.03	0.19	21.1	7.71	-1.2	-0.05	-594.5	0.906	79.2
Kunming	0.40	324.5	-5.41	19.1	-0.89	0.07	12.3	1.65	8.5	14.6	-4627.8	0.952	45.9
Nanjing	0.33	-154.4	2.42	-10.6	0.46	0.10	-19.2	-0.34	-9.1	74.6	2135.5	0.994	22.1
Guangzhou	0.30	103.6	-0.83	14.5	-0.42	0.09	-17.1	-1.80	-8.1	79.3	-2612.3	0.995	17.3

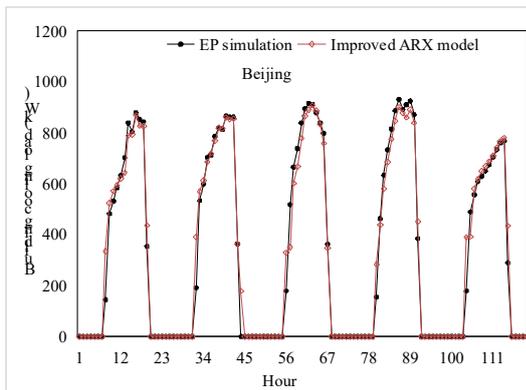**Figure 2.** Building cooling load prediction in Harbin.

Figure 3. Building cooling load prediction in Beijing.
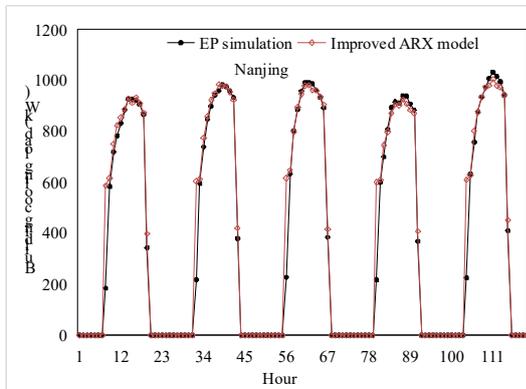

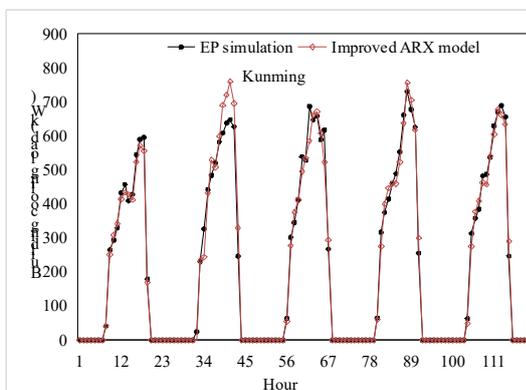**Figure 4.** Building cooling load prediction in Nanjing.

Figure 5. Building cooling load prediction in Kunming.

Figure 6. Building cooling load prediction in Guangzhou.

Table 4. Statistical evaluation of improved ARX model for 5 cities.

City	R^2	RMSE (kW)	EEP (%)	CV (%)	MAE (kW)
Harbin	0.936	43.3	6.7	9.7	16.7
Beijing	0.920	63.3	7.0	9.6	21.1
Kunming	0.942	45.9	6.0	10.3	35.1
Nanjing	0.925	67.1	6.7	8.6	13.1
Guangzhou	0.928	76.3	6.9	9.3	28.5

5 Conclusions

In this paper, an improved ARX model was developed for office buildings cooling load prediction in the five major climates in China, namely severe cold, cold, hot summer and cold winter, temperate, and hot summer and warm winter. This model improved the prediction accuracy of regression analysis methods by including higher order and interaction terms based on traditional ARX model. Higher order and interaction terms are used to manage multi-parameter dynamic changes. The improved ARX model consists of eight variables, with eleven coefficients accessed at a time. The predicted cooling load data was compared with EnergyPlus simulated data of each city. Results showed that the improved model had good predicted accuracy in five climatic zone. The deviations are largely within ±10% according to the relative errors for the validated time span in different cities. The R^2 values were from 0.920 in Beijing to 0.942 in Kunming, indicating a high accuracy of cooling prediction. The prediction results also demonstrated that the improved ARX model can adapt to different climatic conditions, including those nonlinearity cases. The improved model can be used to guide the operation of chiller units and achieve energy savings by predicting the cooling load. The high accuracy and simple mechanism of the improved ARX model make it feasible for practical applications. It may be used for assisting the operation of air-condition systems so as to achieve optimal controls and energy efficiency.

Acknowledgements

References

1. R. Jing, M. Wang, R. Zhang, N. Li, Y. Zhao, Energy Build. **144,** 117-128 (2017).

2. Y. Yao, J. Chen, Energy Build. **42**, 570-583 (2010) .

3. 3.　R. Sarwar, H. Cho, S. Cox, P. Mago, R. Luck, Energy. **119**, 483-496, (2017).

4. G. Qiang, T. Zhe, D. Yan, Z. Neng, Energy and Build. **107**, 445-455 (2015).

5. T. Catalina, V. Iordache, B. Caracaleanu, Energy Build. **57**, 302-312 (2013) .

6. F. Khayatian, L. Sarto, G. Dall'O', Energy Build. **125**, 45-54 (2016).

7. Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida, Energ Convers Manage. **50**, 90-96 (2009) .

8. M. Shin, S. Do, Energy Buildings. **110**, 57-70(2016).

9. A. Aranda, G. Ferreira, M. Mainar-Toledo, S. Scarpellini, E. Llera Sastresa, Energy Build. **49**, 380-387 (2012).

10. Y. Guo, E. Nazarian, J. Ko, K. Rajurkar, Energ Convers Manage. **80**, 46-53 (2014) .

11. J. Lam, K. Wan, D. Liu, C. Tsang, Energ Convers Manage. **51**, 2692-2697 (2010).

12. *Code for design of heating ventilation and air-conditioning* (GB 50019-2003). Beijing: China Planning Press (2003) [in Chinese].

13. R. Henninger, M, *ASHRAE standard* (2013).

14. P. Huang, G. Huang, Y. Wang, Energy Buildings. **91**, 26-36 (2015) .