An hourly electricity consumption calculation method for hvac terminal units with classification and regression tree on the basis of sub-metering

IOP Conf. Series: Earth and Environmental Science **238** (2019) 012002 doi:10.1088/1755-1315/238/1/012002

An hourly electricity consumption calculation method for hvac terminal units with classification and regression tree on the basis of sub-metering

Y Ji[1,2]*, P Xu[3] and J Y Chen[4]

[1]Beijing Key Laboratory of Green Built Environment and Energy Efficient Technology, Beijing University of Technology, Beijing, 100124, China

[2]College of Architecture and Civil Engineering, Beijing University of Technology, Beijing, 100124, China

[3]School of Mechanical Engineering, Tongji University, Shanghai, 201804, China

[4]Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong

jyesse2015@163.com

Abstract. Building energy efficiency is an important section of energy conservation, while the premise of mining energy-saving potential is clearly obtain the specific energy consumption of each building service system. So electricity sub-metering was widely implemented in China in recent years. Main building energy systems and high-power electrical equipment all can be metered directly except the HVAC terminal units (like fan coils and air handing units), because they are always mixed with lighting system caused by existing electricity circuit design. It is an obstacle to get detailed and systematic sub-metering data. Therefore, an indirect calculation algorithm is necessary. Considering the fact that the amount of historical sub-metering data is huge, the algorithm should be calculated fast and ensure a certain degree of accuracy. And given the variety of building types, it is better that the algorithm also has classification function. To solve above problem, an automatic HVAC terminals energy use calculation method is proposed in this paper. The core algorithm of this method is classification and regression tree (CART). In this research, the establishment of CART model is achieved and the selection of input variables is discussed. Finally, this method is demonstrated and validated in an actual building located in Shanghai.

1. Introduction

Building sector energy use accounted for 27.5%, 41.3% and 40% of total energy use in China in 2014 (Building energy conservation research center of Tsinghua University 2017), the United States in 2010 (US Department of Energy 2010) and Europe in 2010 (Council, E.P.A 2010), respectively. So building energy efficiency plays an important role in energy conservation field. And the basis of mining energy-saving potential is clearly obtain the specific energy consumption of each building energy system.

The building total electricity consumption can be metered in four main sub-meters and each sub-meter also can be divided into secondary items (MOHURD 2008, Ji et al. 2015). Nearly all main and secondary sub-meters can be measured directly except the electricity consumption of HVAC terminal units, because the electricity consumption of HVAC terminal units (like fan coils and air handing units) is always mixed with lighting-plug or power sub-meter caused by existing electricity circuit design. Survey results indicate that HVAC terminal units energy use is mixed with lighting-plug sub-meter in most cases. However, it is of great significance to obtain a clear sub-metering data for many fields, such as the pre-assessment of building energy-saving potential, the calculation and verification of energy-saving rate and the fault detection and diagnosis of HVAC system. In view of these, it is necessary to find an indirect way to obtain the energy consumption data of HVAC terminal units.

Pandit, Wu, Seem and Braun et. al. found that the energy consumption of almost all commercial buildings showed periodic characteristics (Pandit and Wu 1983, Seem and Braun 1991). Then Claridge et. al. (1991) did researches and concluded that lighting and equipment energy consumption changed with a day or year cyclical and was free from the impact of outdoor weather parameters. Dhar et al. made further developments and pointed out that the energy use characteristics of working days and non-working days in commercial buildings were very different, so they should be modelled separately (Dhar, Reddy and Claridge 1998, 1999, 1999). Ji and Xu (2015) built hourly Fourier series models to disaggregate the hourly HVAC terminal units energy use from the mixed lighting-plug sub-meter and power meter. The disaggregation results are good, but in this method the classification of the day type is more rough and they did not achieve the automation of computing and the standardization of processes.

Previous studies were carried on using Fourier series models because the algorithm itself has periodic characteristics. Taking it into account that the day type classification of different buildings may be different. So if an algorithm has duality characteristics of classification and regression, it will be well worth adopting. The classification and regression tree (CART) model is exactly in line with this expectation. Hence, in this research work the CART algorithm is attempted to calculate the hourly electricity consumption of HVAC terminal units.

2. Methodology

This section mainly introduces the principle of classification and regression tree algorithm and builds the model to solve the problem raised above.

2.1. Introduction of CART algorithm

Decision tree is a method commonly used in data mining (Lior and Oded 2008.). A tree model includes root node, internal node, leaf node and the criteria for selecting branch attributes which are information gain, entropy and Gini impurity. Tree models where the target variable can take a finite set of values are called classification trees, while the target variable can take continuous values are called regression trees. Classification and regression tree algorithm, also called CART algorithm, was first proposed by Breiman in 1984 (Breiman et al. 1984). It is a specific decision tree algorithm and widely used currently. The criteria for selecting branch attributes in CART algorithm is Gini impurity. The establishment procedure of a CART model is illustrated in Fig. 1.

2.2. Establishment of specific CART algorithm

Existing studies have shown that hourly electricity consumption of lighting-plug in commercial buildings changes periodically and has nothing to do with external factors such as weather parameters. Lighting-plug energy consumption is interruption in the whole year, while the energy use of HVAC systems just appears in heating and cooling seasons. On the basis of the above, a model can be built, using the hourly lighting-plug sub-meter in transition seasons to train the model and predicting the hourly lighting-plug sub-meter in cooling and heating season. The mixed hourly energy use of lighting-plug and HVAC terminal units can be measured in real buildings. So the hourly energy consumption of HVAC terminal units can be calculated by:

$$E_{ter} = E_{mix} - E_{lig} \quad (1)$$

where E_{lig} is the calculated hourly lighting-plug sub-meter in cooling and heating season (kWh), E_{mix} is the mixed hourly energy use of lighting-plug and HVAC terminal units (kWh), and E_{ter} is hourly energy consumption of HVAC terminal units (kWh).

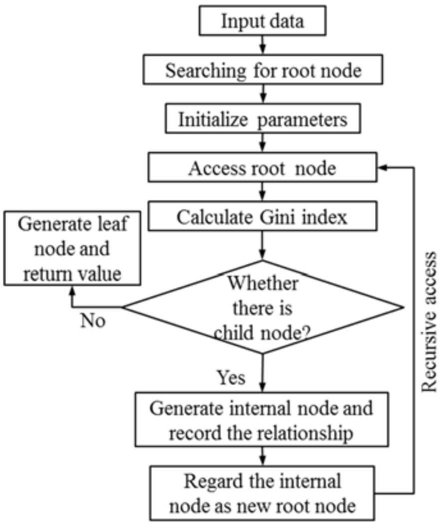


Figure 1. The establishment procedure of a tree model with CART algorithm

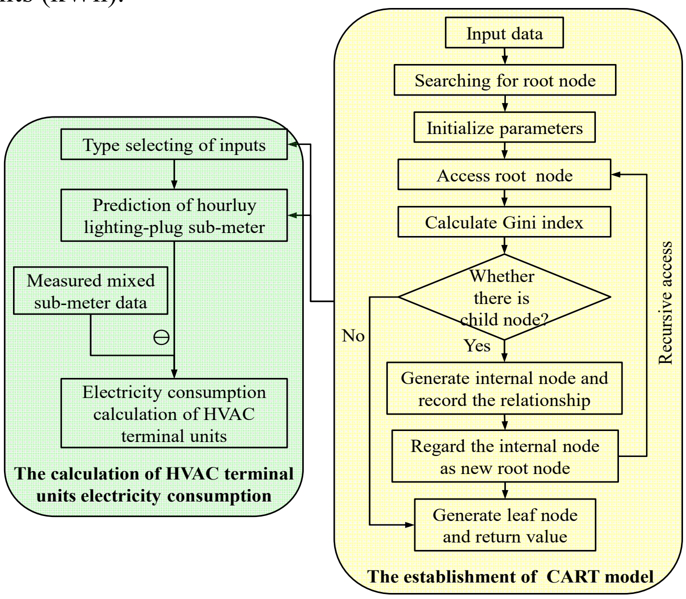


Figure 2. The model for calculating hourly electricity consumption of HVAC terminal units

The procedure of the model for calculating the hourly energy consumption of HVAC terminal units is illustrated in Fig. 2. The procedure is divided into two main parts: (1) the calculation of hourly lighting-plug sub-meter and (2) the calculation of hourly HVAC terminal units electricity consumption.

2.2.1. Determination of the input parameter types

Previous researches showed that hourly electricity consumption of lighting-plug in commercial buildings changes periodically according to time. So the main work in this section is making a time schedule. Analyse the timetable according to the difference of time periods:

✧ Year (Y): 365 or 366 days per year, marked as 1~365(or 366).
✧ Month (M): 12 months per year, marked as 1~12.
✧ Day (D): there are three classification methods:
  • D1: Monday~Sunday (1~7), Holiday (8) and Compensated workday (0), marked as 0~8.
  • D2: Weekday (1), Weekend (0) and Holiday (2), marked as 0~8.
  • D3: Workday (1) and Non-workday (0), marked as 0~1.
✧ Hour (H): there are two classification methods:
  • H1: 24 hours per day, marked as 1~20.
  • H2: The opening time is different in different buildings.

The above time tags can be combined into different timetable input patterns and the input patterns are summarized in Table 1.

Table 1. Summary table of time schedule input patterns

No.	Input patterns	No.	Input patterns
1	Y-M-D1-H1	7	M-D1-H1
2	Y-M-D1-H2	8	D1-H1
3	Y-M-D2-H1	9	D2-H1
4	Y-M-D2-H2	10	D3-H1

5	Y-M-D3-H1	11	H1
6	Y-M-D3-H2		

Aim to each time variable input pattern, analysed the hourly electricity consumption prediction results of the lighting-plug sub-meter. The coefficient of variability (CV, Eq. (2)) and mean relative error (MRE, Eq. (3)) are selected as the evaluation index of the result accuracy. And the multiple-running-average time is taken as the evaluation index of the computing speed. Comprehensively considering the accuracy and speed, the appropriate model type can be selected.

$$CV=\left(\frac{1}{N}\sum_{i=1}^{N}\left(E_{Mi}\text{-}E_{Pi}\right)^{2}\right)^{\frac{1}{2}}\left(\frac{1}{N}\sum_{i=1}^{N}E_{Mi}\right)^{-1} \tag{2}$$

$$MRE=\left(\sum_{i=1}^{N}\left|E_{Mi}\text{-}E_{Pi}\right|\right)\left(\sum_{i=1}^{N}E_{Mi}\right)^{-1} \tag{3}$$

In above equations E_{Mi} is measured electricity energy use data of i^{th} data point (kWh), E_{Pi} is predicted electricity energy use data of i^{th} data point (kWh), N is the total number of data points.

2.2.2. Electricity consumption prediction of lighting-plug sub-meter

Complete the above section, final model for calculating the hourly electricity consumption of lighting-plug sub-meter has been built. Use the final model to predict the hourly electricity consumption of lighting-plug sub-meter in cooling and heating season, and use the CV and MRE index to evaluate the results.

2.2.3. Electricity consumption calculation of HVAC terminal units

After above work, the hourly electricity consumption of lighting-plug sub-meter in cooling and heating season has been predicted. And in the real building the mixed hourly energy use of lighting-plug and HVAC terminal units is measured. The hourly energy consumption of HVAC terminal units can be calculated by Eq.(1). The same as above, use the CV and MRE index to evaluate the results.

3. Case study

3.1. Introduction of the building

The model built above is demonstrated and validated in an actual commercial building located in Shanghai, in which the hourly electricity consumption of HVAC terminal units is measured directly. The researched building information is listed in Table 2.

Table 2. Descriptions of the commercial building researched in this case study

Building name	A
Location	Shanghai
Type of building	Complex building
Floor	Shopping mall(1-4), Office(5-34)
High (m)	150
Area(m^2)	68,330
Opening time	Shopping mall: Full year: 10:00~22:00; Office: Weekday: 8:00~18:00
Type of HVAC terminals	AHU(1-4), FCU+DOAS(5-34)
Type of energy use	Electric

The opening time of every building energy system is confirmed through investigation, listed in Table 2. And the input data types and the times are listed in Table 3.

3.2. Determination of the input parameter types

The time schedule is prepared according to the description in Section 3.1. And the special opening time is adjusted on the basis of Table 3, from 8:00 to 22:00. Aim to every time variable input pattern, use the hourly electricity consumption of the lighting-plug sub-meter in transition season to train the CART models and use those data in cooling and heating seasons to validate the models. The results of building A are listed in Table 4. In this study the multiple-running-average time is the mean time of 1000 runs and presented by $\overline{T}$ (s). The MRE and CV values cover the working period in cooling season, working period in heating season and non-working period.

From Table 4, taking model No. 1&2, No.3&4 and No.5&6 as three pairs, it can be seen that in each pair the calculation accuracy in working time is the same, but in non-working time H1 is better than H2. So H1 is selected between H1 and H2. And in model No. 8, 9 and 10, the calculation accuracy are similar but the computing speed of Model No. 10 is the fastest. So D3 is better among D1, 12 and D3. Model No. 11 is the fastest with 0.0127 s but its accuracy is obviously worse than other models. Its MRE and CV in working time are all larger than 12%, while those of other models are all smaller than 5%. And model No. 1 is the most accurate but its computing speed is too slow compared to other models. It is nearly 14 times as much as that of model No.11. Considering both the speed and the accuracy, model No. 10 is the best for building A, which means this model is affected by day types and hours. Considering both the speed and the accuracy, select two weeks data to train model No. 10.

3.3. Electricity consumption prediction of lighting-plug sub-meter

Complete the work of the above two sections, final model for calculating the hourly electricity use of lighting-plug sub-meter has been built for this researched commercial building. Building A uses the model No. 10 with two weeks training data and the calculation results are in Table 5.

Table 3. Input data selection of the commercial building researched in this case study

lighting-plug sub-meter	electricity consumption of HVAC terminal units	
	Cooling season	Heating season
2013.01.01-2013.12.31	2013.06.01-2013.09.30	2013.01.01-2013.03.31; 2013.12.01-2013.12.31

Table 4. Calculation result of each time variable input pattern of building A

No.	Input pattern	Validation model MRE(%)			Validation model CV(%)			$\overline{T}$ (s)
		Work_cooling	Work_heating	Non-work	Work_cooling	Work_heating	Non-work	
1	Y-M-D1-H1	1.00	1.22	1.89	1.39	1.91	2.75	0.1780
2	Y-M-D1-H2	1.00	1.22	5.20	1.39	1.91	7.08	0.1432
3	Y-M-D2-H1	1.09	1.18	2.82	1.50	1.79	4.58	0.2503
4	Y-M-D2-H2	1.09	1.18	6.09	1.50	1.79	7.94	0.1892
5	Y-M-D3-H1	1.09	1.18	2.88	1.49	1.79	4.58	0.2090
6	Y-M-D3-H2	1.09	1.18	6.13	1.49	1.79	7.95	0.1706
7	M-D1-H1	1.55	1.88	3.38	2.18	2.81	5.06	0.1394
8	D1-H1	2.30	3.34	4.57	2.99	4.60	6.56	0.0195
9	D2-H1	2.38	3.48	5.96	3.16	4.84	8.30	0.0148
10	D3-H1	2.38	3.48	6.25	3.16	4.84	8.65	0.0140
11	H1	12.28	14.02	22.82	15.42	18.23	37.31	0.0130

Table 5. Calculation result of each training data size of building A

Validation model MRE(%)			Validation model CV(%)		
Work_cooling	Work_heating	Non-work	Work_cooling	Work_heating	Non-work
2.46	4.60	6.60	3.30	6.13	9.12

3.4. Electricity consumption calculation of HVAC terminal units

After completing the above work, the hourly electricity use of lighting-plug sub-meter in cooling and heating season has been predicted. The hourly energy consumption of HVAC terminal units is calculated and evaluated in this section. The calculation results of building A is illustrated in Fig. 3. The evaluation index are the MRE, the CV and the data point proportion in different error ranges. The MRE and CV results are summarized in Table 6. The data point proportion in different error ranges are shown in Fig. 4.

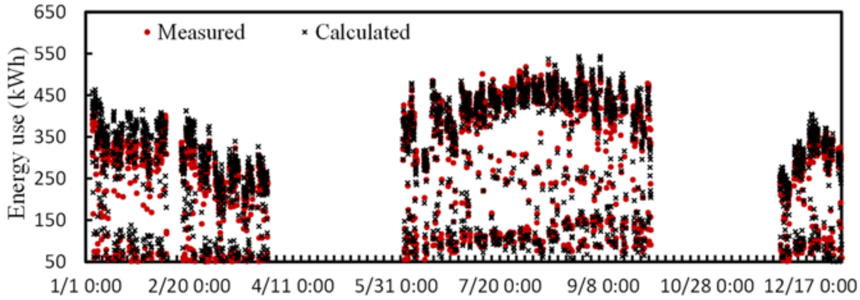


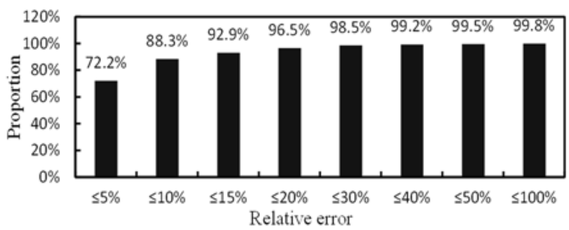
Figure 3. Electricity consumption calculation result of HVAC terminal units

The abscissa axis and vertical axis of Fig. 3 represent the time in hour and the electricity consumption of HVAC terminal units in kWh. The red dots indicate the measured values in building A in working time, and the black cross represents the values in working time calculated by the model built in this research. It can be seen that the calculated data and the measured data are matching greatly.

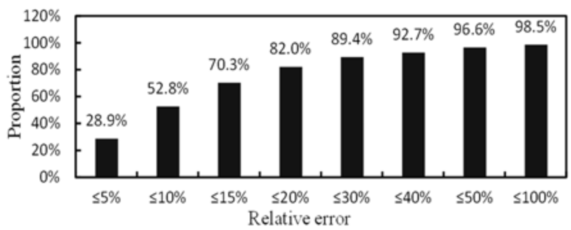
As shown in Table 6, the MRE values of the energy consumption calculation result of HVAC terminal units in building A in working time of cooling season and heating season and non-working time are 3.77%, 10.71% and 23.03%, respectively. And the CV values are 5.07%, 12.52% and 30.98%. Next look at Fig. 4, the abscissa axis represents the relative error of absolute value and vertical axis is the proportion of data points falling within the scope of the relative error requirement in the total data points. The total number of data points calculated in cooling period is 1275 and that in heating period is 1230. Taking the relative error of 20% as example, the proportions of data falling within this range are 96.5% in cooling period and 82.0% in heating period.

Table 6. Electricity consumption calculation result of HVAC terminal units in building A

MRE(%)			CV(%)		
Work_cooling	Work_heating	Non-work	Work_cooling	Work_heating	Non-work
3.77	10.71	23.03	5.07	12.52	30.98



(a) Cooling period



(b) Heating period

Figure 4. The data point proportion in different error ranges of building A

4. Conclusions

Through the study above, some conclusions can be drawn. CART algorithm with dual functions of classification and regression is well suited to solve the problem presented in introduction that calculating the hourly energy use of HVAC terminal units indirectly where they cannot be measured directly. A generic model and standardized process are built, and the type of input variables and the amount of training data are analysed in this research. This method can be fully programmed and is very suitable for processing a large amount of data in different types of buildings. It is worthy to being promoted in the cities where sub-metering is highly implemented. Hope to get weather data, occupancy data and more sub-meter data in buildings that have independent measurement of HVAC terminal units in the future and do some follow-up work.

Acknowledgements
The research is financially supported by General Financial Grant from the China Postdoctoral Science Foundation (2017M620561) and the grant from Beijing Municipal Human Resources and Social Security Bureau (2017ZZ050).

References
[1] Breiman L, Friedman J, Olshen R, et al. 1984. Classification and Regression Trees. Belmont, CA: Wadsworth International Group. Biometrics. Vol.40(3), pp. 17-23.
[2] Building energy conservation research center of Tsinghua University. 2017. Annual Report on China Building Energy Efficiency 2017. https://www.amazon.cn/dp/B072BW7SY8. [In Chinese]
[3] Claridge DE, Haberl JS, Turner D, O'Neal DL. 1991. Heffington WM, Tombari C, et al., Improving energy conservation retrofits with measured savings. ASHRAE Journal. Vol. 33, pp. 10-14.
[4] Council, E.P.A. 2010. Directive 2010/31/EU of the European Parliament and of the Council of 19 May, 2010 on the energy performance of buildings. Official Journal of the European Union. pp. 13–35.
[5] Dhar A, Reddy TA, Claridge DE. 1998. Modeling Hourly Energy Use in Commercial Buildings With Fourier Series Functional Forms. Journal of Solar Energy Engineering. Vol.120(3), pp. 217-223.
[6] Dhar A, Reddy TA, Claridge DE. 1999. A Fourier series model to predict hourly heating and cooling energy use in commercial buildings with outdoor temperature as the only weather variable. Journal of Solar Energy Engineering. Vol. 121(1), pp. 47-53.
[7] Dhar A, Reddy TA, Claridge DE. 1999 Generalization of the Fourier series approach to model hourly energy use in commercial buildings. Journal of Solar Energy Engineering. Vol. 121(1), pp. 54-62.
[8] Lior Rokach，Oded Maimon. 2008. Data mining with decision trees: theory and applications. World Scientific Pub Co Inc. ISBN 978-9812771711.
[9] Ministry of Housing and Urban-Rural Development of the People's Republic of China (MOHURD). 2008. Technical guideline for sub-metering data collection of energy consumption monitoring system in state-owned buildings and large-scale public buildings. Document No. 114 [2008] Annex I. Jun. 24.
[10] Pandit SM, Wu SM. 1983. Time Series and System Analysis with Applications. John Wiley and Sons, New York.
[11] Seem JE, Braun JE. 1991. Adaptive methods for real-time forecasting of building electrical demand. ASHRAE Transactions. 97.
[12] US Department of Energy, 2010. Buildings Energy Data Book.
[13] Ying Ji, Peng Xu, Yunyang Ye. 2015. HVAC terminal hourly end-use disaggregation in commercial buildings with Fourier series model. Energy and Buildings. Vol. 97, pp. 33-46.