Paper ID APEN-MIT-2019_250
Applied Energy Symposium: MIT A+B
May 22-24, 2019 • Boston, USA

A comparative analysis of data-driven building energy benchmarking methods: a case study in China

line 1: 1st Xue Liu
line 2: *Joint International Research Laboratory of Green Buildings and Built Environments (Ministry of Education), Chongqing University*
line 3: *National Centre for International Research of Low-carbon and Green Buildings (Ministry of Science and Technology), Chongqing University*
line 4: Chongqing, China
line 5: cheerps12@126.com

line 1: 2nd Yong Ding
line 2: *Joint International Research Laboratory of Green Buildings and Built Environments (Ministry of Education), Chongqing University*
line 3: *National Centre for International Research of Low-carbon and Green Buildings (Ministry of Science and Technology), Chongqing University*
line 4: Chongqing, China
line 5: dingyongqq@163.com

line 1: 3rd Hao Tang
line 2: *Joint International Research Laboratory of Green Buildings and Built Environments (Ministry of Education), Chongqing University*
line 3: *National Centre for International Research of Low-carbon and Green Buildings (Ministry of Science and Technology), Chongqing University*
line 4: Chongqing, China
line 5: 1157663718@qq.com

Abstract—A reasonable building energy efficiency benchmarking program plays an important role in energy consumption control and supervision. Previous studies have focused on the process of establishing a single benchmarking method, but few have compared the performance and outcomes of different methods. To fill this gap, this paper selects two benchmarking methods—multiple linear regression (MLR) based on Energy Star, and stochastic frontier analysis (SFA) to develop benchmarking models. We demonstrate each method using data on the energy and building characteristics of 45 four- and five-star hotel buildings located in Chongqing, China. To compare the consistency and explanatory ability of two methods, we first utilize the Spearman rank correlation analysis to test whether these methods have consistent energy efficiency ranks and then present Sankey diagrams to further reveal the interactions of the estimated energy efficiency scores obtained from these methods. The results show that even though the ranks of sampled buildings are basically consistent, the energy efficiency scores have significant differences especially for the buildings with low energy efficiency scores. Furthermore, we discuss the explanatory ability of each method. In addition to building characteristics, the design and operational characteristics of the HVAC system have great effects on building energy consumption.

Keywords—energy benchmarking, multiple linear regression, stochastic frontier analysis, comparative analysis

I. INTRODUCTION

The total energy consumption for building operation reached 857 million ton of standard coal equivalent (Tce), taking up 20% of the total energy consumption [1]. Facing the contradiction of an increasing building energy consumption and global warming, resources shortage as well as environmental pollution, Chinese government proposed "Energy Supply and Consumption Revolution Strategy (2016–2030)" in 2017 [2], indicating that china's energy strategy should be transformed gradually from enhancing the efficiency of supply-side to the management of demand-side. Specifically, an implementation of the double control of total quantity and intensity should be promoted in China. Committed to the peak of carbon emission in 2030 [3], it is urgent to propose a regional benchmarking program in order to keep up with the requirements of energy conservation and emission reduction in China.

Building energy benchmarking defines a public yardstick of energy-use performance during a period [4]. Commonly used benchmarking methods can be classified to simulation analysis and data-driven methods. The simulation method more focus on single or several buildings performance evaluation rather than benchmark for city-scale buildings because of lacking detail physical characteristics of buildings. The energy performance indicator or energy usage intensity (EUI) is simple to use for ranking sampled buildings and understanding the energy usage level. While Sharp made an argument that the simple EUI could not reliably represent the level of energy consumption performance due to ignoring other important building physical and operation characteristics [5]. To improve the reliability of a single energy benchmarking indicator and reduce the dimensions of the building characteristics of sampled buildings, multiple linear regression models developed by EUI and energy consumption drivers are widely applied in building performance benchmarking and have already adopted by ENERGY STAR. Despite the ease implementation of ENERGY STAR, the residuals from a regression model not only represent the relative inefficiency of a building, but also statistical noise and unexplained factors. To improve these defects, some studies [6][7] adopted stochastic frontier analysis (SFA) approach to evaluate energy efficiency performance. The advantage of SFA compared with the multiple linear regression (MLR) approach is that this model can estimate inefficiency and data noise from the deviations from the frontier and the actual energy consumption by the assumptions about the distribution of the measurement errors and the inefficiency terms [8]. Another data-driven approach discussed in current studies is the clustering technique, which classifies buildings based on multiple dimensions of building features using the k-means clustering method [9]. Whereas, the main shortcoming of clustering method is that we do not know how to classify the new buildings that are not in the samples we use for clustering.

Although there are various benchmarking methods mentioned in previous studies, few have compared these methods in terms of whether they can provide consistent benchmarking results or which method has the most robust performance. Therefore, this study selected two representative benchmarking methods: MLR based on Energy Star and SFA to build benchmarking models respectively by using a sample of hotel buildings in Chongqing. Then, a comparative analysis of two benchmarking methods is conducted in terms of consistency.

II. Methodology and Materials

A. Data acquisition

Supported by the Chongqing Municipal Commission of Housing and Urban-Rural Development, a total of 48 hotel buildings were collected from energy audit reports. These reports contained monthly energy bills for at least 12 months, including electricity and natural gas, as well as 25 detailed physical and operation information which affecting energy consumption. After removing the buildings with incomplete information, a valid sample of 45 hotel buildings was determined. The statistical summary of partial information of the sample is shown in TABLE I.

TABLE I. STATISTICAL SUMMARY OF PARTIAL COLLECTED DATA OF SURVEYED HOTELS

Values	Gross floor area (m^2)	Occupancy rate (%)	Building stories number	Total energy per unit area (kWh/m^2/year)
Max	97425	0.92	61	318.7
Min	10984	0.30	4	39.8
Average	39078	0.62	19	154.4
Standard deviation	19800	0.16	11.51	65.6

B. Method A: multiple linear regression analysis (MLR)

The concept of method A based on multiple linear regression analysis (MLR) is to obtain a normalized EUI by regression, and then evaluate the building energy consumption level through comparison of the actual energy consumption and normalized energy consumption. Before establishing a regression model, the form of EUI should be determined according to correlation analysis between total energy consumption and various potential drivers. Pearson correlation coefficient was used to test the most significant factor affecting building energy consumption. If the impact on the energy consumption of certain factor is much higher than the others, the unit energy consumption of that factor should be considered as the dependent variable.

In order to determine the normalized EUI, a stepwise multiple linear regression model was developed by introducing significant factors, removing insignificant factors step by step, and then getting a more accurate regression model with no serious multicollinearity. The model can be given by equation (1).

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \varepsilon \tag{1}$$

Where Y is EUI, β_0 is the intercept, β_i are the regression coefficients, X_i represents the significant influencing factors (i=1,2, ..., n), and ε is the random error.

Using the benchmarking model, normalized EUIs for each building can be calculated to reflect the average energy usage level. To evaluate the energy efficiency level of a building, the energy efficiency ratio (EER) should be used, which can be obtained by equation (2).

$$EER = \frac{EUI_{Actual}}{EUI_{predicted}} \tag{2}$$

Based on the result of EER, the cumulative probability density curve was first generated, which describes the relationship of the ratios and cumulative probability density for energy efficiency rankings, then a gamma distribution function was adopted to fit the curve. For a certain ratio, the ENERGY STAR score can be calculated as (1 - cumulative probability) × 100. The higher score means the higher ranking and more efficient for a building.

C. Method B: stochastic frontier analysis (SFA)

The stochastic frontier analysis (SFA) is a parameter frontier approach, which uses regression analysis and a mathematical formula to form the frontier line, assuming the existence of a parametric function between production inputs and outputs [10]. The main advantage of SFA compared with multiple linear regression approach is that it can separate a random error item from an inefficiency item [11]. In this paper, the main objective is to apply SFA to obtain the specific efficiency ratio for each sampled building. Cobb-Douglas cost function was selected as the form of the frontier model, as shown in equation (3).

$$\ln y_i = \beta_0 + \sum_{k=1}^{k} \beta_k \ln x_{ki} + u_i + v_i, i = 1,2,\ldots,n \tag{3}$$

Where y_i represents the actual EUI (calculated by the ratio of annual energy consumption and gross floor area) of a building and i is building numbers; x_k is a vector of potential factors of EUI and k is factor numbers; the β coefficients pertain to the potential factors; u_i and v_i represent inefficiency and random error respectively. The parameters of β, u_i and v_i were estimated by maximum likelihood techniques after some assumptions were given. In particular, the inefficiency item u_i was assumed to be $u_i \sim iidN^+(0, \sigma_u^2)$, and the noise item v_i was assumed to be $v_i \sim iidN(0, \sigma_v^2)$.

By calculating the lower-bound frontier, the minimum cost of EUI can be estimated by some energy influencing factors so that the energy efficiency for each building can be determined. This study only considered two situations containing four and five inputs combinations in SFA model. This is related to the limited amount of sample buildings and various potential independent variables, which may result in overfitting in the model. Enumerate method was used to list all the possible inputs combinations, then the output of SFA model could be calculated based on each combination. The likelihood ratio test was adopted to check whether the linear cost function was rejected in favor of the cost function, and the criteria was the chi-squared distribution with 1 degree of freedom from SFA model. If it is significantly larger than the critical value, then the null-hypothesis can be rejected (the null-hypothesis where there is no difference between the buildings in terms of efficiency with SFA model comparing with ordinary regression model). The final efficiency score of each building can be calculated by this model. The energy efficiency TE_i for each building can be estimated as:

$$TE_i = \frac{y_i}{f(x_i;\beta)e^{v_i}} = \exp(-u_i) \tag{4}$$

The energy efficiency TE_i is defined as the ratio between the minimum energy consumption and the actual consumption and takes values of between 0 and 1. The closer the efficiency measure is to 1, the more efficient the building operation can be considered.

III. RESULTS

A. Method A

1) Determination of EUI

This study used Pearson correlation coefficient to determine annual energy use with the selected 25 potential energy drivers, then we selected the most significant driver as normalized variable. Pearson correlation coefficients of annual energy consumption with the significant drivers were calculated and are shown in TABLE II. GFA is the most significant factor of annual energy consumption and the correlation coefficient is clearly higher than other factors, which conforms to the conclusion in most studies [12][13]. This result indicates that energy use per unit area can be taken as the indicator of evaluating EUI of hotel buildings in this sample.

TABLE II. THE PEARSON CORRELATIONS OF SIGNIFICANT DRIVERS FOR ANNUAL ENERGY CONSUMPTION

Variable name	Description	Pearson correlation	P- value
ZONE	Hotel location district	0.405**	0.006
STAR5	5-star hotel	0.475**	0.001
GFA	Gross floor area	0.774**	0.000
STORY	Number of building stories	0.469**	0.001
ROOM	Number of guest rooms	0.477**	0.001
CONVEN	GFA for convention centers and office area	-0.382**	0.010
CP	Cooling period	0.451**	0.002
HP	Heating period	0.318*	0.033

*Correlation is significant at the 0.05 level (2-tailed).

**Correlation is significant at the 0.01 level (2-tailed).

2) Stepwise regression model

The stepwise regression method was used to obtain the relationship between EUI and independent variables. Three independent variables were reserved from using the stepwise selection method (criteria: probability of F to enter$\leqslant$0.05, probability of F to remove $\geqslant$0.10). TABLE III. provides the results of independent variables in multiple linear regression model. It is found to be significant at the 0.0001 level. The squared multiple correlation coefficient (R2) was 0.700 with adjusted R2=0.677, and thus it can adequately explain the correlation between the EUI and the independent variables.

TABLE III. RESULTS OF MULTIPLE REGRESSION ANALYSIS FOR EUI

Variable	Non-standardized coefficient		Standardized coefficients beta	p-value	overall R^2
	B	Standard error			
(intercept)	-18.662	20.647	-	0.371	0.700 0.677[a]
X_{CP}	0.031	0.014	0.325	0.038	
X_{HP}	0.048	0.016	0.450	0.004	
X_{SA}	161.863	62.699	0.236	0.014	

Therefore, the final model can be written as follows:

$$EUI_{normalized} = -18.662 + 0.031X_{CP} + 0.048X_{HP} + 161.863X_{SA} \tag{5}$$

Where X_{CP} and X_{HP} stands for cooling and heating period, respectively; X_{SA} is the total area of entertainment and fitness club provided in hotels.

3) Developing benchmark table

Using the benchmarking model, the predicted EUIs were obtained, then the energy efficiency can be calculated. The gamma curve was fitted according to the efficiency ratio and cumulative distribution, as shown in Fig. 1.

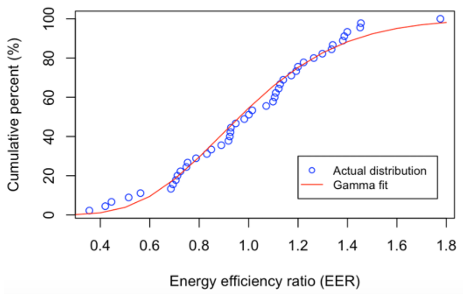

Fig. 1. Cumulative distribution curve of the EER

According to the gamma function, a representative benchmarking rating table can be generated, shown in TABLE IV. For example, if the score is 80, the percentile should be 20%, and the corresponding EER is 0.717.

TABLE IV. BENCHMARKING RATING TABLE

Percentile	EER	Score

10%	0.607	90
20%	0.717	80
30%	0.804	70
40%	0.884	60
50%	0.963	50
60%	1.047	40
70%	1.142	30
80%	1.261	20
90%	1.438	10

B. Method B

The stochastic frontier analysis was performed by using R programming language with 'frontier' package [14]. All available energy drivers were used to determine the most optimal input variables by using Enumerate method. The likelihood ratio test (LR test) was performed to select the best model. The maximum likelihood estimate results of coefficients and parameters for SFA model are shown in TABLE V. and TABLE VI.

TABLE V. COEFFICIENTS ESTIMATE FOR THE SFA MODEL BY THE MAXIMUM LIKELIHOOD TECHNIQUE

Variable	Coefficient value	Standard error	Probability
Intercept	2.104	0.717	0.003***
GFA_{SER}	0.222	0.073	0.002***
FLOOR	0.161	0.091	0.074*
HTC_{EWALL}	0.104	0.127	0.412
HVAC1	0.348	0.107	0.001***

TABLE VI. PARAMETER ESTIMATES FOR THE SFA MODEL BY THE MAXIMUM LIKELIHOOD TECHNIQUE

Parameter	Estimate
λ	3.678
σ_u^2	0.348
σ_v^2	0.025
γ	0.931
$\log[L(H_0)]$	-20.381
$\log[L(H_1)]$	-19.017
LR test	0.049*

The parameter gamma (γ) is close to 1 and shows that the inefficiency item accounts for 93% of the error item, which represents the most variation in the actual EUI that comes from inefficiency, with only 7% of the random error in the error item. An LR test was used to test whether there is a difference between sampled buildings in terms of energy efficiency. According to TABLE VI. , the LR test provides a statistic of 2.73, which exceeds the 5% critical value of 2.71. Hence, this result demonstrates there are significantly differences in efficiency between buildings.

The coefficients of the SFA model, including GFA_{SER}, FLOOR, $HTCE_{WALL}$ and HVAC1, are all positive and show a positive relationship between the corresponding variable and EUI. This result is found to be partly the same as the findings in [6]; that is, a growth in entertainment area in a building will lead to an increase in the building's EUI.

As shown in Fig. 2, the purple line represents the predicted frontier (the minimum EUI with the highest energy efficiency) and is equivalent to the actual EUI, demonstrating that the energy performance of the buildings on the line is efficient. The scatters are closer to the purple line, meaning that the corresponding building is more efficient in its energy utilization. However, several buildings are found to have an actual EUI lower than the frontier because of random error items. Overall, there is a clear gap between the actual EUI and the fitted frontier for sampled buildings according to the slope of the two lines in Fig. 2.

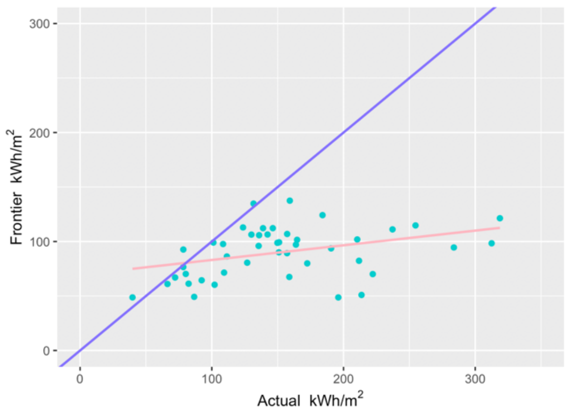

Fig. 2. Relation between predicted frontier and actual EUI

C. Comparison analysis between method A and B

Evaluating the accuracy of each benchmarking method is a challenge for a benchmarking program since it is time consuming and costly to acquire the practical energy efficiency for individual buildings on large-scale sampled buildings, leading to an absence of ground truth to verify the obtained energy efficiency scores of each method. Hence, this section verifies the energy performance by comparing whether the benchmarking results of the methods are consistent. In other words, we focus on the consistency of the benchmarking results instead of the comparison to ground truth of energy performance. In addition, benchmarking does not merely put forward an accurate energy performance assessment or a ranking table; a better explanatory ability is another important property of benchmarking methods. Spearman correlation coefficient was selected to test the association between two ranked variables, which is a nonparametric statistic without critical requirements for the distribution of variables. Spearman correlation coefficients with corresponding p-values in the brackets are shown in TABLE VII.

TABLE VII. SPEARMAN CORRECORRELATION COEFFICIENTS OF RANKING RESULTS

Variables	Energy Star	SFA
Energy Star	1	0.439(0.003)
SFA	0.439(0.003)	1

The result reflects the efficiency scores from two benchmarking methods have significant positive correlation with each other, which indicates the result is consistent. To further detect the differences among the benchmarking methods, we first divided the benchmarking results into several grades. For instance, the results based on MLR and SFA were divided into four grades, similar to the building energy grades proposed in the law passed by New York City [15], which include "0-19", "20-49", "50-89"and "90 or above"; "0-0.19", "0.2-0.49", "0.5-0.89" and "0.9 or above".

Fig. 3 uses Sankey diagrams to present comparison of the benchmarking results based on the two methods. There are 42% of the sampled buildings received different grades. The significant difference mainly exists in low performance; i.e., the clusters of "0-19" and "20-49" from MLR are split in the cluster of "0.5-0.89" from SFA.

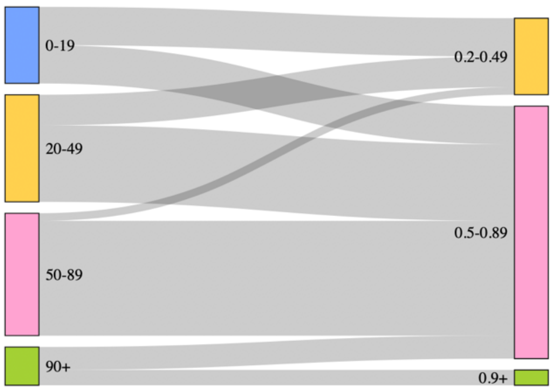

Fig. 3. Comparison of the benchmarking results based on two methods (Left node: MLR, right node: SFA)

In terms of explanatory ability, MLR method based on Energy Star provides three independent variables, namely, the percentage of entertainment and leisure area and cooling and heating periods, that significantly affect the EUIs in the sampled buildings. The variables of cooling and heating periods are related to the operational characteristics of the HVAC system. This result suggests that hotels in this sample should adopt measures that can reduce the operation time of HVAC systems to improve energy performance. The SFA model in this paper indicates that almost all of the input variables are related to design parameters that seldom change during the building operation period, such as the number of building floors, the type of chiller units and the U-values of external walls. The MLR and SFA methods reveal that the determinations of EUI are related to design and operational characteristics of HVAC system, which suggests that energy retrofitting for these sampled buildings could be dependent on effective measurements implemented in HVAC systems.

Although the MLR and SFA benchmarking methods enable provide some potential suggestions for energy conservation measurements for the sampled buildings, they are unlikely to provide detailed solutions for individual buildings where there is high energy use or energy waste. Furthermore, the obtained energy efficiency scores can only reflect the relative energy performance among the sampled buildings. Therefore, on-site energy audits are still needed to further detect the energy saving potential for individual buildings. The best use case for these benchmarking methods is to determine which buildings are considered to be inefficient in a sample and try to understand the core drivers of energy consumption in the buildings. The explanatory ability of the benchmarking results allows different parties, from buildings users and energy service companies to governmental policy-makers, to discover deeper insights that can be used to make better energy efficiency decisions and policies.

IV. Conclusion

Data-driven methods are widely used to benchmark the energy performance of buildings, but few studies have compared the performance or accuracy of these methods. In this paper, we selected MLR based on Energy Star and SFA based on the energy consumption standard in China to establish benchmarking models. The energy and building characteristic information of 45 four- and five-star hotels in Chongqing was collected for a case study. Since it is difficult to obtain the actual energy efficiency of individual buildings, we focused on the consistency instead of the ground truth of the benchmarking methods to quantify their performance via Spearman correlation analysis and Sankey diagrams. The results of Spearman correlation analysis showed that the ranks of the sampled buildings obtained from the two benchmarking methods were basically consistent. However, the estimated energy efficiency scores were found to be greatly different especially for the buildings low energy performance according to the Sankey diagram. This finding confirmed that using only one benchmarking method may lead to distorted evaluations of the energy performance in sampled buildings. Finally, the explanatory ability of benchmarking methods indicates the influence of the design and operational characteristics of the HVAC system on building energy consumption. These characteristics are suggested to be further collected for developing local or national databases and energy performance disclosure policies. In future work, we intend to expand the sample size by using the public database, like CBECS in the U.S. to test the consistency of energy performance assessment based on these benchmarking methods.

Acknowledgment

This work was supported by the project "Standard System for Post-assessment of Green Building Performance" (the National Key Projects in 13th Five-Year, projected No. 2016TFC0700105).

References

[1] China Association of Building Energy Efficiency. Research report on energy consumption in China, 2017.

[2] X. Y. Yang, S. C. Zhang,W. Xu, "Impact of zero energy buildings on medium-to-long term building energy consumption in China," Energy Policy. Beijing, vol.129, pp. 574–586, June 2019.

[3] U. S. Government, China Government, U.S.-China Joint Announcement on Climate Change, November 2014, available at: https://obamawhitehouse.archives.gov/the-press-office/2014/11/11/us-china-joint-announcement-climate-change

[4] W. Chung, "Review of building energy-use performance benchmarking methodologies," Applied Energy. Hongkong, vol.88, pp.1470-1479, May 2011.

[5] T. Sharp, "Energy Benchmarking In Commercial Office Buildings," Summer Study Proceedings. California. United States, vol.4, pp.321-329,1996.

[6] J. Buck, D. Young, The potential for energy efficiency gains in the Canadian commercial building sector: A stochastic frontier study, Energy. 32 (2007) 1769–1780.

[7] Z. Yang, J. Roth, R.K. Jain, DUE-B : Data-driven urban energy benchmarking of buildings using recursive partitioning and stochastic frontier analysis, Energy Build. 163 (2018) 58–69.

[8] D. Aigner, K. Lovell, P. Schmidt, Upland Development and Prospects for the Rural Poor: Experience in Northern Thailand, J. Econom. 6 (1977) 21–37.

[9] X. Gao, A. Malkawi, A new methodology for building energy performance benchmarking: An approach based on intelligent clustering algorithm, Energy Build. 84 (2014) 607–616.

[10] K. Cullinane, T. F. Wang, D. W. Song, P. Jing, "The technical efficiency of container ports: Comparing data envelopment analysis and stochastic frontier analysis", Transportation Research Part A. Newcastle, vol.40, pp.354-374, May 2006.

[11] D. Aigner, C. A. K. Lovell, P. Schmidt, "Formulation and estimation of stochastic frontier production function models", Journal of Econometrics. U.S.A, vol.6, pp.21-37, July 1977.

[12] J. Wang, "A study on the energy performance of hotel buildings in Taiwan," Energy and Buildings. Taiwan, vol.49, pp. 268–275, June 2012.

[13] Y. J. Xin, S. L. Lu, N. Zhu, W. Wu,"Energy consumption quota of four and five star luxury hotel buildings in Hainan province, China", Energy and Buildings, Tianjin, vol.45, pp.250-256, February 2012.

[14] T. Coelli, A. Henningsen, R Package "frontier", 2017 (Accessed on 17 March 2017) .

[15] Institute for Market Transformation. Coming to NYC: Building Energy Grades; 2018. < https://www.imt.org/coming-to-nyc-building-energy-grades/ > (Accessed 24 June 2019).