Journal of Cleaner Production 338 (2022) 130615

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

Data-driven assessment of room air conditioner efficiency for saving energy

Weiqi Wang [a,1], Zixuan Zhou [b,1], Zhongming Lu [b,*]

[a] Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong Special Administrative Region

[b] Division of Environment and Sustainability, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong Special Administrative Region

ARTICLE INFO

Handling Editor: Mingzhou Jin

Keywords:
Smart meter
XGBoost
Shapley additive explanations
Energy efficiency
Room air conditioner

ABSTRACT

Room air conditioners (RACs) are one of the high energy-consuming home appliances. Developing a smart solution to evaluate and track the efficiency of RACs is essentially useful for residents to make necessary maintenance or replacement decisions. While smart meters are increasingly installed to monitor electricity use, the application of smart meter data to evaluate the RAC efficiency remains inadequate. In this paper, we present a data-driven framework to identify non-inverter window RACs with low energy efficiency by analyzing smart meter data from the university student hall rooms. In this framework, we first applied the extreme gradient boosting (XGBoost) method to predict a RAC's hourly electricity consumption. Then we measured the effect of outdoor temperature on the XGBoost prediction of hourly RAC electricity consumption using the Shapley Additive Explanation method to interpret the RAC's efficiency. We conjectured that the RAC efficiency is normal if the predicted hourly electricity consumption is significantly correlated with the outdoor temperature. In contrast, the RAC efficiency is low if the outdoor temperature changes have little impact on the predicted electricity consumption. Finally, we applied the K-Means clustering algorithm to separate the RACs into the "low efficiency" and "normal efficiency" categories based on each's pattern of outdoor temperature's impact on electricity consumption. Our cross-validation result showed that the XGBoost model can achieve an average R^2 score of 0.50 and an average root mean squared error of 0.20 kWh. We used RAC replacement records to validate our framework of interpreting the RAC's efficiency. On average, RACs having low efficiency consumed 25.69% more electricity per hour. Overall, our data-driven framework can contribute to extending the value of smart meters for RAC efficiency evaluation. Meanwhile, the smart meter data-driven framework can be improved, and more validation is needed in the future.

1. Introduction

The energy demand for space cooling has tripled over the past thirty years, making air conditioners the fastest-growing energy use appliance worldwide. In 2019, about 8.5% of world electricity use came from air conditioning, resulting in nearly 1 Gigatonne of CO_2 emissions (IEA, 2020). Air conditioners are expected to dominate the energy demand growth in buildings by 2050 and become one of the major CO_2 emitters that contributes to global climate change (IEA, 2018). In Hong Kong, room air conditioners (RACs) are responsible for 38% of the total electricity consumption of the residential sector (Electrical and Mechanical Services Department Hong Kong Government, 2020). RACs will continue dominating Hong Kong residential electricity consumption, making the RACs a hotspot for reducing energy consumption and greenhouse gas emissions.

A mandatory energy efficiency labeling scheme has been implemented in Hong Kong to inform customers about energy efficiency when selecting home appliances, including the RACs. However, there is little information regarding the energy efficiency of the RACs during the use phase to facilitate the necessary maintenance or replacement for saving energy. Replacement of inefficient RACs in time can effectively avoid energy wastage and reduce energy consumption (Alabdulkarem and Almutairi, 2020). But it is challenging for residents to estimate and track the efficiency of the RACs at home. Measuring the RAC efficiency can be very technical (Yang et al., 2020), so it is implausible for residents to measure the RAC efficiency regularly. Usually, residents only consider replacing the RAC when there are significant deficiencies (e.g., noise and refrigerant leakage). Likely, the efficiency has already dropped before

* Corresponding author.
E-mail address: zhongminglu@ust.hk (Z. Lu).

[1] Contributed Equally.

https://doi.org/10.1016/j.jclepro.2022.130615
Received 24 August 2021; Received in revised form 7 December 2021; Accepted 16 January 2022
Available online 19 January 2022
0959-6526/© 2022 Elsevier Ltd. All rights reserved.

the significant deficiencies occur, and the electricity has been wasted. Without any measurement tool to retrieve RAC efficiency information, residents can be reluctant to replace RACs because of the cost. Therefore, developing a smart solution for residents to track the RAC efficiency is essentially useful for maintaining high energy efficiency and saving energy.

More real-time electricity data is available as smart meters are being widely installed. Various researchers have investigated the application of smart meter data. Past studies have primarily focused on predictive building energy consumption modeling and explored how to improve the prediction accuracy using data-driven machine learning algorithms (Amasyali and El-Gohary, 2018; Candanedo et al., 2017; Sobrino et al., 2019). These studies are valuable to support the application of demand-side response (Qi et al., 2020; Satre-Meloy et al., 2020), the optimal real-time dispatch (Anvari-Moghaddam et al., 2016), and the increase of renewable energy penetration (Dato et al., 2020; Pawar et al., 2020). Other applications of smart meter data include the classification of electricity load profiles (Piscitelli et al., 2019) and the detection of the existence and usage of RACs in residential buildings (Chen et al., 2020; Liang and Ma, 2020). These studies highlight the potential of generating valuable insights from smart meter data to understand energy-related behaviors. Given the increasing interest in making a decision beyond prediction from smart meter data, it is meaningful to explore the possibility of applying smart meters to evaluate RAC efficiency (Susan, 2017).

However, few studies have discussed the application of smart meter data to evaluate energy efficiency. A common practice to benchmark energy efficiency is to compare energy performance within similar groups (Ahmed Gassar et al., 2019). Power utilities provide the comparison results between the customer and his/her neighborhood to inform him/her of the energy use efficiency. Iyengar et al. developed a data-driven approach named WattScale to identify energy-inefficient buildings and possible causes of inefficiency (Iyengar et al., 2021). WattScale first defines the equations that explain building energy consumption (i.e., heating, cooling, and baseload) and uses Bayesian inference to estimate the equations and parameters (i.e., constant and slopes). The judgment of building efficiency is based on the estimated constants and slopes that describe the energy consumption pattern of buildings. In general, the methodology of applying smart meter data to evaluate the RAC efficiency remains underexplored.

In summary, our study was motivated by 1) the critical need to develop a smart solution for residents to track the RAC efficiency and 2) the increasing interest in going beyond prediction to support decision-making with smart meter data. We hypothesize that smart meter data can be applied to evaluate the RAC efficiency, which can unlock more value of smart meters. To validate this hypothesis, we aim to develop a smart meter data-driven framework to evaluate the efficiency of non-inverter window normal RACs. Our study objectives include 1) developing a smart meter data-driven framework using interpretable machining learning methods to evaluate the RAC efficiency, 2) validating the performance of our data-driven framework by comparing model evaluation with replacement records; and 3) quantifying the electricity saved by recognizing RACs with low energy efficiency. Our smart meter data-driven framework contributes to understanding the feasibility of expanding the value of smart meter data for evaluating the RAC's efficiency. Our framework can be valuable for future studies to validate and improve the smart meter data-driven efficiency assessment, which can help manage the RACs smartly and save energy.

The rest of our paper consists of four more sections. In section 2, we describe our data collection, details of the data-driven framework, and result validation. Results are reported in Section 3 including the outputs from each step of the data-driven framework, the performance of our data-driven framework, and the quantification of energy saving by recognizing RACs with low efficiency. Discussions are presented in Section 4 followed by the conclusion in Section 5.

2. Materials and methods

Fig. 1 summarizes different stages of our research method development, including data collection, framework development, and result validation.

2.1. Data

We collected hourly electricity consumption of non-inverter window normal RACs from smart meter installed in one student hall of our university. Window normal RACs are widely used in residential buildings in Hong Kong for their low prices and easy installation. Fig. 2 shows the appearance of this type of RAC and the external walls of the on-campus student hall from which we collected the smart meter data. We selected our study case due to the data accessibility, reliability, and coverage. Data privacy was one of the critical barriers to access smart data.

The smart meter data were from April 3rd to October 3rd, 2019, in which period the RACs in the building were used frequently. We removed the hours when RACs were turned off. Moreover, RACs might be turned on before the end of an hour, so only a small amount of electricity consumption was recorded in that respective hour. To minimize such bias, for each operating period of any air conditioners, we excluded the data at the beginning hour and the ending hour so that only the hours when RACs were fully functional were considered in our models. In addition, we only used the data when the temperature was in the range of 22–33.5°C so that extreme temperatures in the afternoon were excluded since only a few students stayed in halls at that time. These exclusions made our interpretation analysis more rigorous and convictive. A total of 125,323 data points for 129 rooms were finalized to develop our smart meter data-driven framework.

We collected hourly meteorological data, including the outdoor temperature, precipitation, relative humidity, and irradiance, from the Atmospheric & Environmental Real-time Database (https://envf.ust.hk/dataview/gts/current/). We collected hourly Wi-Fi connection counts to approximate the number of people in the room.

2.2. The smart meter data-driven framework for evaluating the RAC efficiency

Our framework can be divided into three steps (Fig. 1): model the RAC's electricity consumption, interpret the RAC's efficiency, and cluster RACs by efficiency. In the first step, we selected the Extreme Gradient Boosting (XGBoost) algorithm to build our prediction model of hourly RAC electricity consumption. In the second step, we applied the Shapley Additive Explanations to measure the effect of outdoor temperature on the hourly electricity prediction to interpret the RAC's efficiency. Finally, we used the K-Means algorithm to cluster RACs into two categories based on the Shapley value of outdoor temperature.

2.2.1. XGBoost-based hourly RAC electricity consumption modeling

XGBoost algorithm was selected to develop a data-driven hourly RAC electricity consumption predictive model for each of the 129 rooms. The model was trained on the hourly meteorological data and the Wi-Fi data of each room to predict the hourly RAC electricity consumption. XGBoost is an optimized distributed gradient boosting library designed to be highly efficient, flexible, and portable (Chen and Guestrin, 2016). The core algorithm for XGBoost to make predictions is gradient boosting. Gradient boosting is a tree-ensemble technique where new models are created to predict the residuals or errors of prior models. Then the new models are assembled to make the final prediction. It uses a gradient descent algorithm to minimize the loss when adding new models. Multiple regression trees are constructed and assembled sequentially based on the input data by fitting a parameterized function (base learner) to current residuals with least-squares at each iteration (Friedman, 2002). We chose XGBoost because it is highly dominative

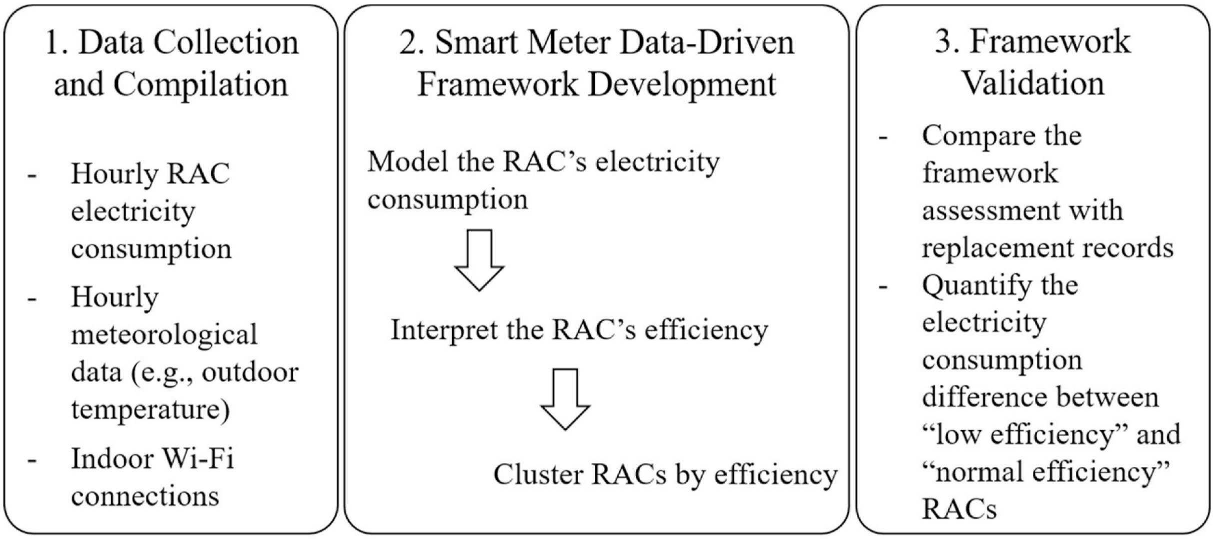

Fig. 1. A flowchart summarizing different stages of our research method development.

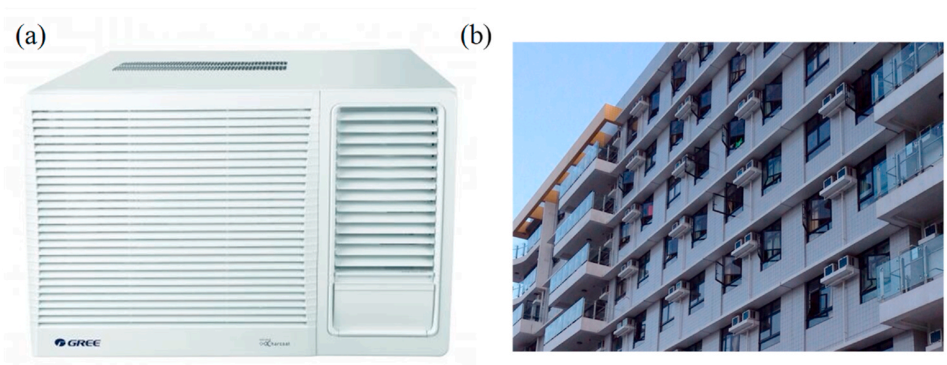

Fig. 2. (a) The typical non-inverter window normal air conditioner sold in Hong Kong; (b) the external walls of the university student hall with window ACs.

and efficient on structured or tabular datasets to perform classification and regression predictive tasks. And more importantly, the algorithm is highly explainable compared to deep learning techniques such as neural networks or Long-Short-Term-Memory (LSTM) as XGBoost inherits the advantages of the regression trees. Precisely, K decision tree models are constructed to output the results according to the input x_i, and the final prediction value is the sum of the K output results (Eq. 1).

$$\widehat{y}_i = \varphi_{x_i} = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathscr{F} \tag{Eq. 1}$$

where $\mathscr{F} = \{f(x) = w_{q(x)}\}(q : \mathbb{R}^m \rightarrow T, w \in \mathbb{R}^T)$ is a set of all possible Classification And Regression Trees (CART); q represents a binary tree which splits the features continuously and assigns data points to the corresponding leaf nodes; f_k represents the function of the k_{th} regression tree; w is a vector of the scores of leaf nodes; T is the number of leaf nodes; and K is the number of the CART trees.

The objective function of XGBoost consists of two parts (Eq. 2). The first part $\sum_{i=1}^{n} l(y_i, \widehat{y}_i)$ is the loss function that measures the difference between the predicted and the observed RAC electricity consumption. In our task, we selected the squared loss (Eq. 3) as the loss function because squared loss is effective in regression tasks (Wang et al., 2017). The second part $\sum_{k=1}^{K} \Omega(\mathrm{f_k})$ is the regularization term. The regularization term also contains γ and λ as regulating parameters for the number of leaf nodes and the scores of leaf nodes to prevent overfitting.

$$Obj = \sum_{i=1}^{n} l(y_i, \widehat{y}_i) + \sum_{k=1}^{K} \Omega(\mathrm{f_k}) \ \ where \ \ \Omega(f) = \gamma \mathrm{T} + \frac{\lambda||\mathrm{w}||^2}{2} \tag{Eq. 2}$$

$$\sum_{i=1}^{n} l(y_i, \widehat{y}_i) = \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 \tag{Eq. 3}$$

For each room, the XGBoost prediction model was asked to minimize the objective function based on the input data.

Table 1 summarizes the statistics of input features after feature engineering (Zhang et al., 2018). Apart from the hourly meteorological conditions and Wi-Fi connection counts, we also included the total electricity use in the previous 1, 3, and 5 h and whether the RAC was turned on during the previous 1 and 2 h. These additional features reflected the impact of the previous hour's operation on the current hourly electricity consumption.

The second measure implemented was Synthetic Minority Oversampling Technique (SMOTE) to balance the distribution of the hourly RAC electricity consumption (Chawla et al., 2002). Fig. 3a demonstrates that a higher portion of hourly RAC electricity consumption data is lower than 0.7 kWh, which can be explained by the operation mode of the window normal RAC. When the room temperature reaches the thermostat, the RAC unit will be turned off while the fan continues to run for the normal window RAC without the inverter. The hourly electricity consumption is small when the RAC unit is turned off. Data-balancing algorithms are required since unbalanced data is fatal to

Table 1
Definitions and statistics of input features for the XGBoost model to predict the hourly electricity consumption of RACs.

Input Features	Description	Min	Max	Mean	Std
Outdoor temperature	Hourly outdoor temperature (°C)	17.39	37.21	26.59	2.92
Humidity	Hourly outdoor relative humidity (%)	9.32	95.25	65.63	19.42
Irradiance	Hourly outdoor irradiance (W/m^2)	0.0	909.03	120.23	206.00
Precipitation	Hourly precipitation (mm)	0.0	43.43	0.33	1.80
Wi-Fi connection counts	Average hourly Wi-Fi connection count to approximate the number of people in the room	0.0	15.1	2.05	1.54
Prev_1hr_AC	The total electricity consumption of the RAC in the previous 1 h	0.0	1.98	0.48	0.34
Prev_3hrs_AC	The total electricity consumption of the RAC in the previous 3 h	0.0	5.71	1.26	0.94
Prev_5hrs_AC	The total electricity consumption of the RAC in the previous 5 h	0.0	9.49	1.91	1.46
Prev_1hr	Whether the RAC was turned on in the previous 1 h (True/False)	0	1	0.85	0.36
Prev_2hrs	Whether the RAC was turned on in the previous 2 h (True/False)	0	1	0.72	0.45

supervised tree algorithms, including XGBoost (Cieslak and Chawla, 2008). To implement the SMOTE, we first divided the dataset into two groups with hourly electricity consumption higher and lower than 0.7 kWh, respectively. SMOTE resampled each group until the sample size of the two groups was balanced. Fig. 3 shows the aggregated distribution of the hourly RAC electricity consumption of all rooms before and after the SMOTE algorithm.

The third measure was hyperparameter tuning, a common practice to improve the prediction accuracy of machine learning models. We selected the Distributed Asynchronous Hyper-Parameter Optimization (Hyperopt) to find the optimal hyperparameters (Bergstra et al., 2015). Hyperopt uses Bayesian optimization for parameter tuning, allowing the user to obtain the best parameters for a given model within a large parameter space at high speed. After hyperparameter tuning, we used the 10-folder cross-validation technique to train and test the XGBoost models for estimating hourly RAC electricity consumption in individual rooms. To measure the prediction accuracy, we used the R^2 score and root mean squared error (RMSE) as the cross-validation evaluation metrics.

2.2.2. *Interpreting the RAC efficiency with Shapley value*

In this study, RAC efficiency was defined as the amount of electricity consumed to provide a certain amount of cooling. A RAC with low efficiency consumes more electricity to provide the same amount of cooling as a normally efficient RAC. Outdoor temperature is a major factor that affects the cooling demand and the consequent electricity consumption (Wu et al., 2017). We conjectured that for a normally efficient RAC, a low outdoor temperature can reduce the predicted hourly RAC electricity consumption to below mean, while a high outdoor temperature can increase the predicted hourly RAC electricity consumption to above mean. In contrast, the outdoor temperature variation does not affect the predicted hourly electricity consumption for a RAC with low efficiency.

We selected the Shapley Additive Explanation to measure the effect of outdoor temperature on the hourly RAC electricity consumption learned by the XGBoost model. The Shapley value of one feature is the weighted average of the difference between estimation with and without the input feature of interest for all possible combinations of the remaining input features (Lundberg et al., 2020; Lundberg and Lee, 2017) (Eq. 4).

$$\varphi_j = \sum_{S\subseteq\{x_1,x_2,x_3,\ldots,x_p\}\setminus\{x_j\}} \frac{|S|!(p-|S|-1)!}{p!}\left(f_{X_s\cup\{x_j\}}\left(X_s\cup\{x_j\}\right)-f_{X_s}(X_s)\right) \tag{Eq. 4}$$

where φ_j is the Shapley value of the j^{th} input feature; $\{x_1, x_2, x_3, \ldots, x_p\}$ is the set of all input features; p is the number of all input features; $\{x_1, x_2, x_3, \ldots, x_p\}\setminus\{x_j\}$ is a set of all features without x_j; S is the subset of $\{x_1, x_2, x_3, \ldots, x_p\}\setminus\{x_j\}$, $f_{X_s}(X_s)$ is the prediction of feature subset S; and $f_{X_s\cup\{x_j\}}(X_s\cup\{x_j\})$ is the prediction of feature subset S with x_j.

The Shapley value for a single feature shows how much this input feature pushes the model prediction electricity consumption away from the average predicted RAC electricity consumption for one room. When the Shapley value of an input feature is negative, it means that the inclusion of the input feature reduces the predicted RAC electricity consumption. When the Shapley value of an input feature is positive, it means that the inclusion of the input feature increases the predicted RAC electricity consumption. Fig. 4 illustrates the Shapley values of the input features in a data sample of a room. Features whose inclusion resulted in higher prediction electricity consumption are shown in red, and those resulted in lower prediction electricity consumption are shown in blue. For example, the outdoor temperature of 18.8 °C reduced the predicted hourly RAC electricity consumption. This is reasonable as the cooling demand is low at the outdoor temperature of 18.8 °C.

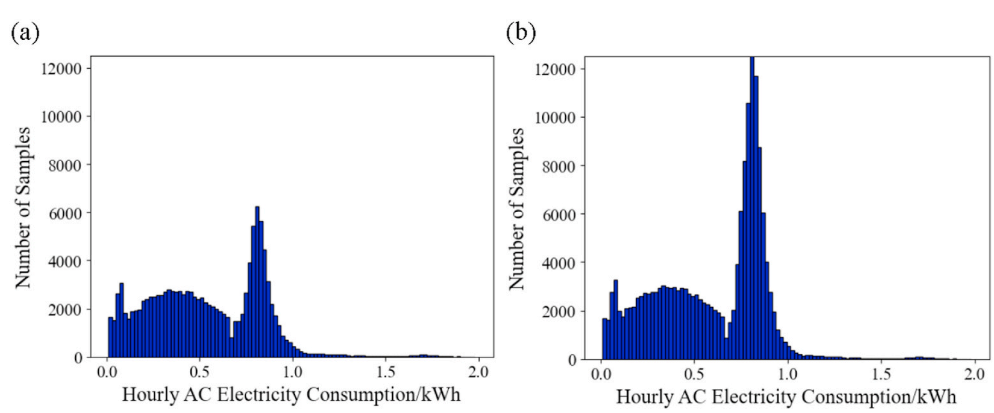

Fig. 3. (a) The distribution of hourly RAC electricity consumption before SMOTE; (b) the distribution of hourly RAC electricity consumption after SMOTE.

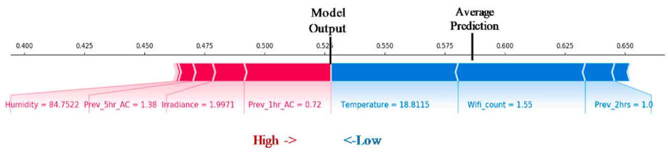

Fig. 4. Shapley values of input features: the blue variable value decreases the prediction while the red variable value increases the prediction.

A python package *SHAP v0.37.0*[2] has been used to calculate the Shapley value of outdoor temperature for the samples of each room, which was then plotted to examine how the change of outdoor temperature affected the prediction of hourly RAC electricity consumption. From these plots, we could observe the distribution pattern of temperature's Shapley value, which visualized the effect of outdoor temperature on predicted electricity consumption. The plots reflected the general trend of how the effect changed as temperature increased, which were used as evidence for assessing the RAC efficiency.

2.2.3. K-means clustering based on RAC efficiency

The Shapley value plots were then clustered into two categories using the K-Means algorithm (Hartigan and Wong, 1979). K-Means clustering is a vector quantization method that aims to partition n observations into a given number of clusters such that the chosen centroid (μ_j) of each cluster has minimized the inertia within that cluster (Eq. 5). Inertia can be recognized as a measure of how internally coherent clusters are.

$$I_C = \sum_{i=0}^{n} \min_{\mu_j \in C}\left(x_i - \mu_j{}^2\right) \tag{Eq. 5}$$

To classify Shapley value images, we flattened each image from a pixel matrix to a vector and then applied the K-Means algorithm to partition 129 vectors into two clusters. Shapley plots within the same cluster shared a similar pattern of how temperature was affecting the electricity consumption of RACs.

2.3. Framework validation

RAC replacement records were collected from the Campus Management Office to help validate our framework. Four rooms replaced their ACs after September 2019, nearly when we finished collecting data. Having confirmed that these RACs are replaced due to general mechanical issues, we considered most of these RACs' efficiency low. We validated our conjecture by examining the Shapely value of the outdoor temperature of these four RACs. Moreover, 27 rooms replaced the RACs from 2016 to March 2019, and there were 93 rooms whose RACs were replaced before 2016. We considered that the percentage of RACs with low efficiency should be smaller among those replaced recently. As another validation, we examined whether our analysis of RAC efficiency presents a similar trend. Afterward, we quantified the potential energy saving if we can identify the RACs with low efficiency to highlight the value of our framework in facilitating smart RAC management.

3. Results

3.1. XGBoost accuracy in predicting the hourly RAC electricity consumption

Fig. 5 summarizes the R^2 score and RMSE of XGBoost models in 10-folds cross-validation when predicting the hourly RAC electricity consumption among the 129 rooms. The mean R^2 score was 0.496, and the RMSE was 0.196 kWh among all the models, indicating that there were still rooms to improve the accuracy. Fig. 6 highlights the comparison between observations and predictions of three rooms with R^2 score of around 0.5, 0.7, 0.9. We found that our prediction could be significantly higher than the observation when the actual hourly RAC electricity consumption was low. This bias can be attributed to the use of Wi-Fi connection counts to approximate the presence of people in the rooms. The approximation can be false when people turn off the RAC and leave the room, but electronic equipment (e.g., laptops, tablets, and smartphones) remains connected. In this case, our XGBoost model could yield a higher predicted hourly RAC electricity consumption considering the presence of people in the room. We also found that the predicted hourly RAC electricity consumption can be much lower than the observation. User behaviors can partially explain this, including leaving the rooms without turning off the RAC, eating foods in the rooms, and returning to the rooms after physical exercises. These behaviors may not happen regularly but do impact the hourly RAC electricity consumption.

3.2. Shapley value of outdoor temperature in predicting the hourly RAC electricity consumption

Fig. 7 shows the Shapley value of different outdoor temperatures in selected rooms. In Fig. 7a, the Shapley value of outdoor temperature increases as the outdoor temperature rises. A negative Shapley value of outdoor temperature means the predicted hourly RAC electricity consumption at the corresponding outdoor temperature is lower than the average prediction among all data samples of one room. On the other hand, a positive Shapley value of outdoor temperature means the predicted hourly RAC electricity consumption at the corresponding outdoor temperature is higher than the average prediction. Fig. 7b demonstrates a very different pattern from Fig. 7a. In Fig. 7b, the outdoor temperature does not influence the prediction of the hourly RAC electricity consumption. Fig. 7 also highlights that the Wi-Fi connections do not influence outdoor temperature's effect on predicted hourly RAC electricity consumption. We set up our conjecture about assessing RAC efficiency based on the response of Shapley value to the outdoor temperature change in predicting the hourly RAC electricity consumption. The RACs presented in Fig. 7a are normally efficient in delivering the cooling service. The RACs in Fig. 7b are considered as having low efficiency.

With 129 Shapley value plots, we applied the K-Means clustering algorithm to cluster these plots into two categories. After noticing significantly false classifications in the normal efficiency category, manual corrections were also carried out, and 5 RACs were moved into the low efficiency group. Finally, we classified 68 RACs into the "normal efficiency" category and the other 61 RACs into the "low efficiency" category.

3.3. Validation of our framework for evaluating the RAC's efficiency

Four rooms replaced the window RACs in September and October 2019, whose Shapley value plots are presented in Fig. 8 (Room A, B, C, and D). In Room A and Room B, the effect of outdoor temperature on the predicted hourly RAC electricity consumption was insignificant. Temperature's impact was also insignificant in Room C when the outdoor temperature ranges between 26 °C and 31 °C. Because there were only a few observation points above 31 °C, as residents do not often stay in the room under that temperature condition, we ignored the trend after 31 °C. In Room D, the increase of outdoor temperature led to a higher predicted hourly RAC electricity consumption. However, this increasing slope was comparatively flat. Overall, the four replaced RACs at the end of our study period presented an insignificant or weak response to outdoor temperature, which aligns with our conjecture in Section 2.2.2.

Fig. 9 shows the number of RACs by replacement year and efficiency classification. The proportion of RACs categorized as "low efficiency" decreases as the replacement time approaches 2019. This is consistent with the fact that the overall performance of recently replaced RACs should be better than that of RACs which have been running for a longer

[2] https://shap.readthedocs.io/en/latest/index.html.

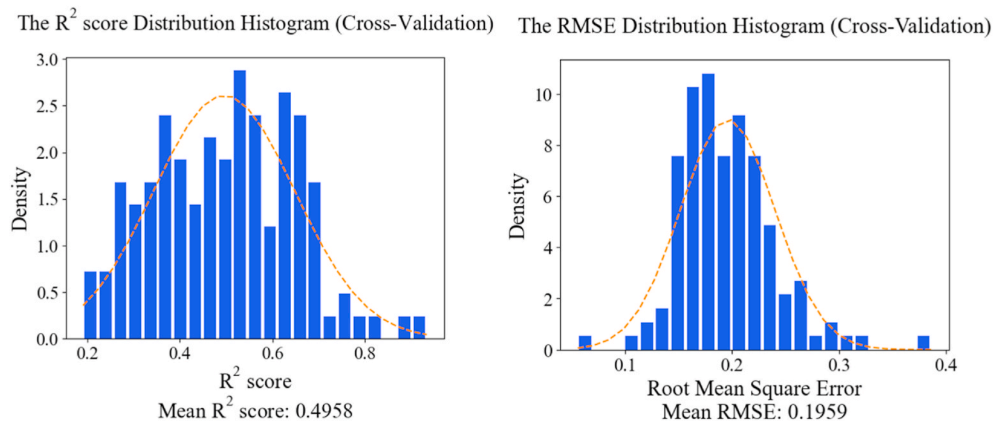

Fig. 5. A summary of the R^2 score and root mean squared error (RMSE) among 129 rooms in 10-folds cross-validation.

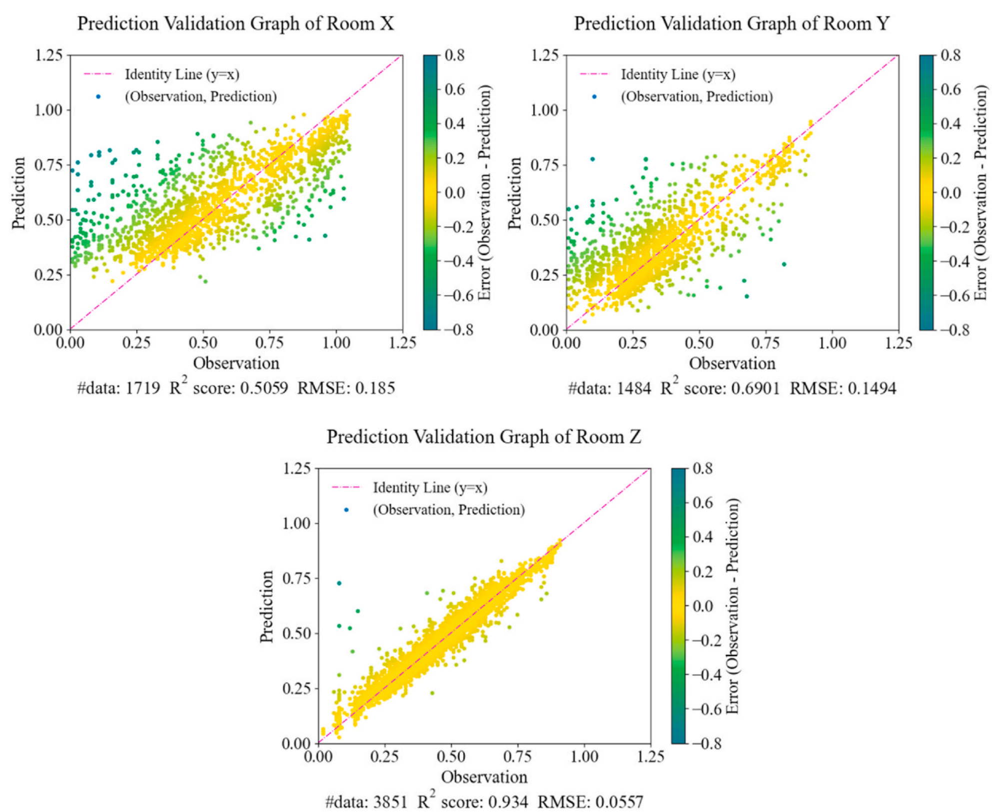

Fig. 6. The comparison between the observations and the predicted hourly RAC electricity consumption among three levels of R^2 score.

time. Meanwhile, our analysis highlighted that around 50% of RACs replaced before or in 2015 are rated "low efficiency" and 11 out of 20 RACs replaced in 2016 and 2017 have low efficiency. It indicated that the year of the installation was not the determining factor of energy efficiency. Our method can provide a good starting point to assess the efficiency of RACs more rationally.

Fig. 10 shows the hourly RAC electricity consumption distribution of the two RAC categories. On average, RACs with low efficiency consume 25.69% more electricity per hour than normal RACs. RACs with low efficiency should be replaced to stop the electricity wastage. In general, these findings proved the practicality of our data-driven framework to assess the RAC efficiency in terms of saving energy and facilitating RAC management.

4. Discussion

In our framework, we selected the XGBoost for both its predictability and interpretability. We can use our XGBoost model to predict the RAC electricity consumption in the next hour. Input features, including the outdoor temperature, humidity, irradiance, and precipitation, are available from weather prediction. Input features regarding RAC use in previous hours can also be calculated accordingly. Input feature of Wi-Fi-connection counts can be estimated based on historical data during the same period. With the input information, the RAC electricity consumption in the next hour can be predicted, which helps occupants and building managers better understand the electrical load of the RAC use. Meanwhile, our framework offers insights about RAC efficiency by measuring the effect of outdoor temperature on the prediction using the Shapley Additive Explanation. Both occupants and building managers

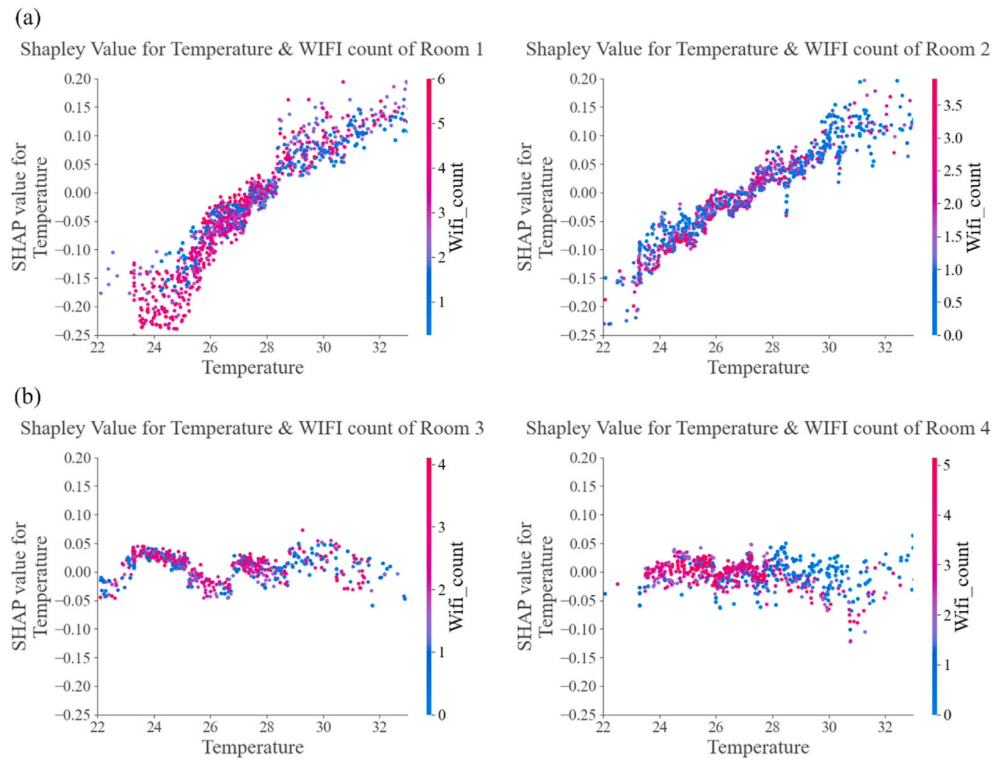

Fig. 7. The interactive Shapley values of outdoor temperatures and Wi-Fi Count for selected rooms: (a) the increase in outdoor temperature increases the prediction; (b) the increase in outdoor temperature does not affect the prediction.

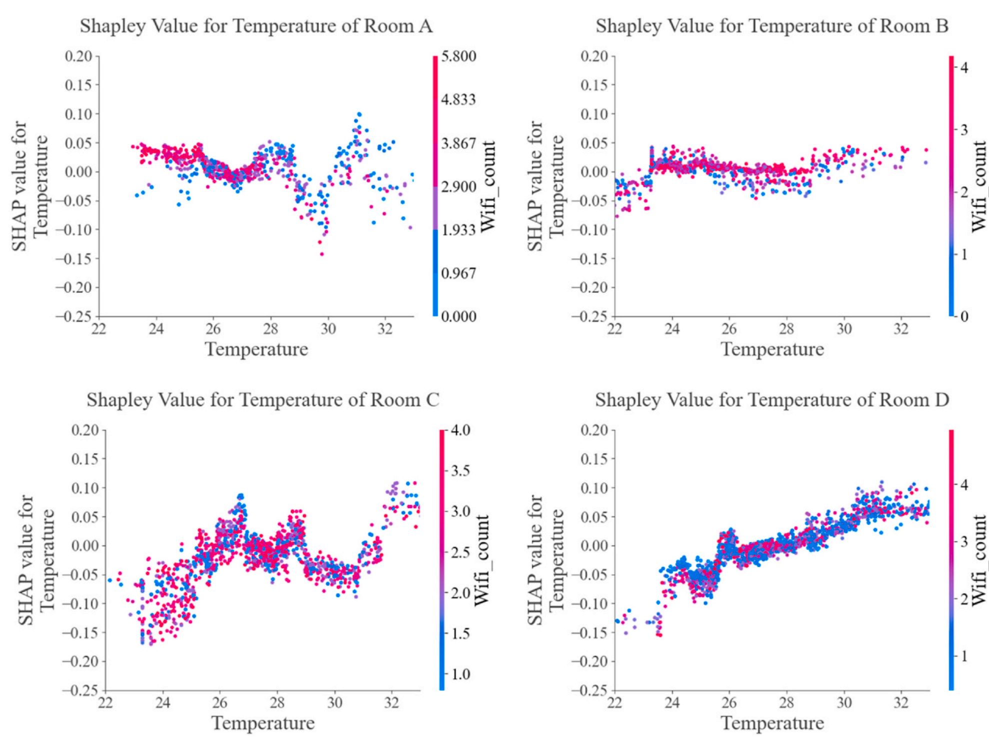

Fig. 8. Shapley value – Temperature/Wi-Fi count plot for four rooms in the “low efficiency” category.

can assess the efficiency of RACs to make informed decisions about repairing or replacing the RACs. Even though we trained and validated our framework based on the window normal RAC, our framework can be extended to study inverter and split type ACs in the rooms. That said, our framework requires more testing and validation in different buildings and regions.

Nevertheless, the accuracy of our prediction model can be further improved. The improvement can be achieved by including additional

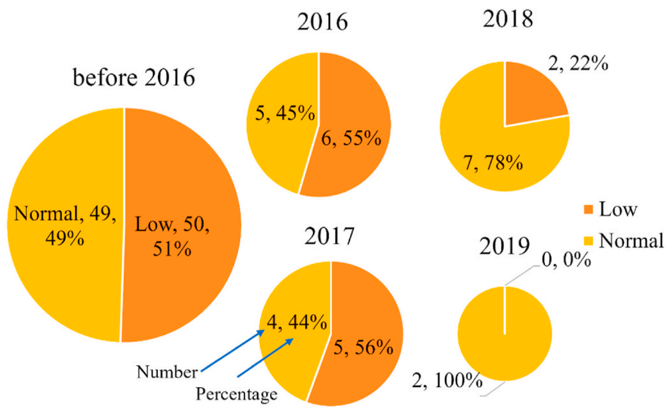

Fig. 9. Classification of RACs by replacement year and operating efficiency.

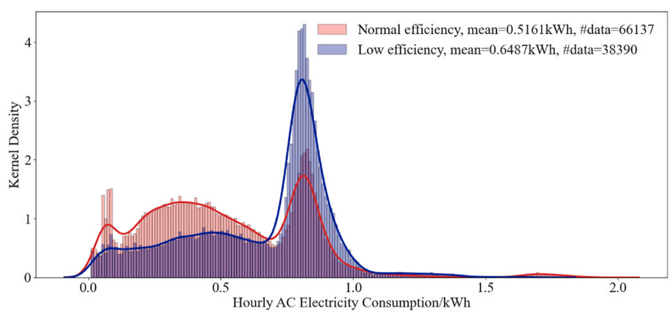

Fig. 10. The distribution of hourly RAC electricity consumption at the two levels of energy efficiency.

features. For instance, the RAC operation setting is one critical piece of information that should be included when predicting the hourly RAC electricity consumption. The RAC operation setting includes the thermostat and fan speed. The thermostat determines whether the RAC unit should be turned on by the difference between the actual room temperature and the pre-setting one. The fan speed can affect how fast the air cycles to cool down the room. These settings of the window normal RAC are adjusted manually. Usually, users may leave the setting of window normal RACs' thermostat default instead of changing it very often. But in some cases, users also adjust the fan speed to cool down the room. People might select a high fan speed if the room is hot and then turn down the fan speed once the room temperature drops. The inclusion of these operation settings can further improve prediction accuracy. However, acquiring these settings' information on an hourly basis is not economically viable for a window normal RAC, which is already cheap. Future tests can be applied to more expensive inverter split type RAC to improve the prediction accuracy. Our prediction model can be improved as well by higher temporal resolution metering data to eliminate the bias caused by turning on/off RACs within 1-h resolution.

To enhance the framework accuracy, our clustering technique can also be optimized in two directions. When preprocessing the Shapley value plots, meaningful dots only take up a small proportion of the entire image, leaving most spaces blank. In this case, stretching the image matrix to a vector will affect the clustering performance as the information brought by colored dots is weakened. To address this issue, autoencoder (Vincent et al., 2008) or self-supervised learning (Van Gansbeke et al., 2020) can be used to extract the features in each plot so that the dot distribution pattern can be better captured for clustering. Additionally, supervised learning such as CNN (Dong et al., 2016) has long-proven success in image classification for its advanced efficiency and accuracy. The RAC efficiency can be measured to label the corresponding Shapley value plots, and a supervised learning network can be trained accordingly. Both directions are expected to improve the image categorization performance to set up a more convictive hypothesis for detecting RACs with low efficiency.

Another direction of future improvement is to explore other interpretable machine learning or deep learning algorithms that can accurately predict the hourly RAC electricity consumption and assess RAC efficiency. XGBoost is a tree-based algorithm. Future studies can explore deep learning algorithms such as deep neural networks (Fekri et al., 2021) and long short-term memory (LSTM) (Wang et al., 2019). The LSTM is a deep learning algorithm for time-series data such as smart meters to make predictions. Deep learning algorithms have demonstrated good performances in prediction, and Shapley value can be estimated to interpret how input features influence the prediction. These future works can also validate our findings using the Shapley outdoor temperature value to assess RAC efficiency.

5. Conclusion

Energy efficiency is critical to saving energy and reducing the impact of energy use. In this study, we developed a novel smart meter data-driven framework to assess the energy efficiency of non-invertor window normal RACs with interpretable machine learning methods. First, we built up an XGBoost prediction model to predict the hourly electricity consumed by RACs. Then we applied the Shapley Additive Explanation method to measure the effect of outdoor temperature on the hourly RAC electricity consumption predicted by XGBoost to interpret the RAC efficiency. Finally, we applied the K-Means algorithm to classify the efficiency of RACs into two clusters based on how outdoor temperature affected the predicted hourly RAC electricity consumption. We validated our framework using the RAC replacement records from the campus facility management office. Our result showed that RACs with low efficiency consumed 25.69% more electricity per hour. Nevertheless, we also acknowledged our limitation in capturing the influence of users' behavioral factors and other additional factors that affect energy consumption such as RAC operation state, which may be one future improvement area. Another direction of future work is to explore other machine learning algorithms for assessing the efficiency of RACs. Overall, our data-driven framework provides an inspiring and reliable starting point for smart management of RACs to save energy.

Reproducibility

Sample data and code can be found at https://github.com/MighTy-Weaver/Inefficient-AC-detection for training XGBoost model and plotting Shapley value of outdoor temperature.

CRediT authorship contribution statement

Weiqi Wang: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, contributed equally. **Zixuan Zhou:** Methodology, Investigation, Resources, Data curation, Writing – original draft, contributed equally. **Zhongming Lu:** Conceptualization, Supervision, Writing – review & editing, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This project was supported by the Undergraduate Research Opportunity Program (UROP) of The Hong Kong University of Science and Technology (HKUST) and the Sustainable Smart Campus project of HKUST. The authors would also like to thank the anonymous reviewers for their valuable comments and suggestions. The views and ideas expressed here belong solely to the authors and not to the funding agencies.

References

Ahmed Gassar, A.A., Yun, G.Y., Kim, S., 2019. Data-driven approach to prediction of residential energy consumption at urban scales in London. Energy 187, 115973. https://doi.org/10.1016/j.energy.2019.115973.

Alabdulkarem, A., Almutairi, Z., 2020. Testing and simulation of potential energy savings in renewing used split-type and window-type air-conditioners. Build. Serv. Eng. Res. Technol. 41, 586–602. https://doi.org/10.1177/0143624419894438.

Amasyali, K., El-Gohary, N.M., 2018. A review of data-driven building energy consumption prediction studies. Renew. Sustain. Energy Rev. https://doi.org/10.1016/j.rser.2017.04.095.

Anvari-Moghaddam, A., Guerrero, J.M., Rahimi-Kian, A., Mirian, M.S., 2016. Optimal real-time dispatch for integrated energy systems: an ontology-based multi-agent approach. In: 2016 IEEE 7th International Symposium on Power Electronics for Distributed Generation Systems (PEDG), pp. 1–7. https://doi.org/10.1109/PEDG.2016.7526997.

Bergstra, J., Komer, B., Eliasmith, C., Yamins, D., Cox, D., 2015. Hyperopt: a Python library for model selection and hyperparameter optimization. Comput. Sci. Discov. 8, 14008. https://doi.org/10.1088/1749-4699/8/1/014008.

Candanedo, L.M., Feldheim, V., Deramaix, D., 2017. Data driven prediction models of energy use of appliances in a low-energy house. Energy Build 140, 81–97. https://doi.org/10.1016/j.enbuild.2017.01.083.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res. 16, 321–357.

Chen, M., Ban-Weiss, G.A., Sanders, K.T., 2020. Utilizing smart-meter data to project impacts of urban warming on residential electricity use for vulnerable populations in Southern California. Environ. Res. Lett. 15, 64001. https://doi.org/10.1088/1748-9326/ab6fbe.

Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery, pp. 785–794. https://doi.org/10.1145/2939672.2939785.

Cieslak, D.A., Chawla, N.V., 2008. Learning decision trees for unbalanced data BT - machine learning and knowledge discovery in databases. In: Daelemans, W., Goethals, B., Morik, K. (Eds.), Machine Learning and Knowledge Discovery in Databases. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 241–256. https://doi.org/10.1007/978-3-540-87479-9_34.

Dato, P., Durmaz, T., Pommeret, A., 2020. Smart grids and renewable electricity generation by households. Energy Econ 86, 104511. https://doi.org/10.1016/j.eneco.2019.104511.

Dong, C., Loy, C.C., He, K., Tang, X., 2016. Image super-resolution using deep convolutional networks. IEEE Trans. Pattern Anal. Mach. Intell. 38, 295–307. https://doi.org/10.1109/TPAMI.2015.2439281.

Electrical and Mechanical Services Department Hong Kong Government, 2020. No Title. Hong Kong.

Fekri, M.N., Patel, H., Grolinger, K., Sharma, V., 2021. Deep learning for load forecasting with smart meter data: online Adaptive Recurrent Neural Network. Appl. Energy 282, 116177. https://doi.org/10.1016/j.apenergy.2020.116177.

Friedman, J.H., 2002. Stochastic gradient boosting. Comput. Stat. Data Anal. 38, 367–378.

Hartigan, J.A., Wong, M.A., 1979. Algorithm as 136: a K-means clustering algorithm. J. R. Stat. Soc. Ser. C (Applied Stat. 28, 100–108. https://doi.org/10.2307/2346830.

IEA, 2020. Cooling. Paris.

IEA, 2018. The Future of Cooling. Paris.

Iyengar, S., Lee, S., Irwin, D., Shenoy, P., Weil, B., 2021. WattScale. ACM/IMS Trans. Data Sci. 2, 1–25. https://doi.org/10.1145/3406961.

Liang, H., Ma, J., 2020. Separation of residential space cooling usage from smart meter data. IEEE Trans. Smart Grid 11, 3107–3118. https://doi.org/10.1109/TSG.2020.2965958.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2, 56–67. https://doi.org/10.1038/s42256-019-0138-9.

Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, pp. 4766–4775.

Pawar, P., TarunKumar, M., Vittal, K.P., 2020. An IoT based Intelligent Smart Energy Management System with accurate forecasting and load strategy for renewable generation. Meas. J. Int. Meas. Confed. 152, 107187. https://doi.org/10.1016/j.measurement.2019.107187.

Piscitelli, M.S., Brandi, S., Capozzoli, A., 2019. Recognition and classification of typical load profiles in buildings with non-intrusive learning approach. Appl. Energy 255, 113727. https://doi.org/10.1016/j.apenergy.2019.113727.

Qi, N., Cheng, L., Xu, H., Wu, K., Li, X.L., Wang, Y., Liu, R., 2020. Smart meter data-driven evaluation of operational demand response potential of residential air conditioning loads. Appl. Energy 279, 115708. https://doi.org/10.1016/j.apenergy.2020.115708.

Satre-Meloy, A., Diakonova, M., Grünewald, P., 2020. Cluster analysis and prediction of residential peak demand profiles using occupant activity data. Appl. Energy 260, 114246. https://doi.org/10.1016/j.apenergy.2019.114246.

Sobrino, E.M., Santiago, A.V., González, A.M., 2019. Forecasting the Electricity Hourly Consumption of Residential Consumers with Smart Meters Using Machine Learning Algorithms. 2019 IEEE Milan PowerTech, pp. 1–6. https://doi.org/10.1109/PTC.2019.8810902.

Susan, A., 2017. Beyond prediction: using big data for policy problems. Science 355, 483–485. https://doi.org/10.1126/science.aal4321 (80-.).

Van Gansbeke, W., Vandenhende, S., Georgoulis, S., Proesmans, M., Van Gool, L., 2020. SCAN: learning to classify images without labels. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 268–285. https://doi.org/10.1007/978-3-030-58607-2_16.

Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P.A., 2008. Extracting and composing robust features with denoising autoencoders. In: Proceedings of the 25th International Conference on Machine Learning, pp. 1096–1103. https://doi.org/10.1145/1390156.1390294.

Wang, S., Dong, P., Tian, Y., 2017. A novel method of statistical line loss estimation for distribution feeders based on feeder cluster and modified XGBoost. Energies 10, 2067. https://doi.org/10.3390/en10122067.

Wang, Y., Gan, D., Sun, M., Zhang, N., Lu, Z., Kang, C., 2019. Probabilistic individual load forecasting using pinball loss guided LSTM. Appl. Energy 235, 10–20. https://doi.org/10.1016/j.apenergy.2018.10.078.

Wu, J., Liu, C., Li, H., Ouyang, D., Cheng, J., Wang, Y., You, S., 2017. Residential air-conditioner usage in China and efficiency standardization. Energy 119, 1036–1046. https://doi.org/10.1016/j.energy.2016.11.038.

Yang, Z., Ding, L., Xiao, H., Zhang, G., Wang, B., Shi, W., 2020. All-condition measuring methods for field performance of room air conditioner. Appl. Therm. Eng. 180, 115887. https://doi.org/10.1016/j.applthermaleng.2020.115887.

Zhang, C., Cao, L., Romagnoli, A., 2018. On the feature engineering of building energy data mining. Sustain. Cities Soc. 39, 508–518. https://doi.org/10.1016/j.scs.2018.02.016.