Energy Proceedings
Vol 60, 2025
ISSN 2004-2965

An Attention-Enhanced Deep Learning Method Tailored for Non-Intrusive Load Monitoring in Air Conditioning Systems#

Chuanyu Tang [1], Yi Gong [1], Qiong Chen [2], Nan Li [1,3,*]

1 Department of Civil Engineering, Chongqing University, Chongqing 400044, China

2 School of Civil Engineering and Architecture, Wuhan University of Technology, Wuhan, Hubei, 430070, China

3 National Center for International Research of Low-Carbon and Green Buildings, Ministry of Science and Technology, Chongqing University, Chongqing 400044, China

(Corresponding Author: nanlicqu@126.com)

This is a paper for the 5th Applied Energy Symposium and Forum: Renewable Energy Matrix (REM2025), Oct. 29-31, 2025, Yancheng, China.

ABSTRACT

The increasing energy demand for air conditioning systems presents substantial economic and environmental challenges, making effective air conditioning load monitoring crucial. Traditional intrusive load monitoring techniques were costly and challenging to implement, whereas non-intrusive load monitoring (NILM) technology offered a cost-effective alternative. This study constructed the CN-AC-NILM dataset, customized for the characteristics of Chinese electrical appliances and cultural usage patterns, and subsequently developed an RNN-attention model for air conditioning systems, with a focus on evaluating the impact of the attention mechanism on the model's performance. Initially, the optimal input feature set for the air conditioning load disaggregation prediction model was determined using the forward feature selection method based on the dataset, followed by an evaluation of performance differences among various RNN models. Subsequently, it was observed that in the comparative analysis of baseline models applied to air conditioning system for NILM, LSTM exhibited higher accuracy than GRU, and bidirectional models outperformed unidirectional models. Furthermore, after introducing the attention mechanism in the baseline models, the average accuracy of RNN models increased by 21.09%, accompanied by a 5.33% increase in average training time. Finally, the effects of attention mechanisms and bidirectional structure on the models were compared, revealing that the introduction of the attention mechanism outperformed the bidirectional structure, and the LSTM-Attention exhibited the best overall performance for NILM of the air conditioning systems. Notably, Attention mechanisms enhanced model interpretability through the visualization of attention weights. By improving load disaggregation accuracy and interpretability, this attention-enhanced model enables more precise energy management for air conditioning systems. It reduces electricity costs, optimizes energy use, and supports demand-side management, offering substantial economic savings and lowering environmental impact, especially in high energy-demand settings.

Keywords: attention mechanism, Recurrent neural n etwork, Non-intrusive load monitoring, Air condition ing system, CN-AC-NILM dataset

NONMENCLATURE

Abbreviations

BiLSTM	Bidirectional Long Short-Term Memory
DT	Decision Trees
GRU	Gated Recurrent Unit
MLP	Multilayer Perceptron
NILM	Non-Intrusive Load Monitoring
R^2	R-squared
SVM	Support Vector Machines
BiGRU	Bidirectional Gated Recurrent Unit
DNN	Deep Neural Network
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
RNN	Recurrent Neural Network
RMSE	Root Mean Squared Error

1. INTRODUCTION

The construction industry globally was the predominant sector in energy consumption. Specifically, the energy consumption of buildings constitutes over 40% of the total global energy, translating to more than 30% of the worldwide CO2 emissions[1,2]. Among the various factors contributing to the substantial energy consumption in buildings, air conditioning systems played a significant role, exerting a notable impact on overall energy consumption[3].

A practical approach to reduce air conditioning energy consumption was through load monitoring[4]. Concurrently, load monitoring of air conditioning systems was also advantageous for enhancing the stability of the electrical grid[5]. However, traditional intrusive methods of electricity load monitoring, while capable of obtaining accurate data, were limited by the requirement to install sensors and other equipment, leading to high costs and significant deployment challenges, which restricted their applicability in small-scale implementations[6,7]. To address this issue, Hart pioneered research on non-intrusive load monitoring (NILM) in the 1980s[8]. In contrast to intrusive load monitoring, NILM required the installation of a sensor at the power outlet only, enabling detailed information about each appliance to be obtained through feature analysis of aggregated signals or specific decomposition algorithms.

Over the past decade, numerous studies have been conducted in the field of NILM. However, due to limitations in datasets quality and algorithm performance, achieving high accuracy in load monitoring has been challenging. With the remarkable achievements of deep learning in the fields of computer vision and natural language processing, there has been a rapid proliferation of Deep Neural Network (DNN) methods to address the NILM problem since 2015. In particular, RNN, as a specialized type of DNN, including Long Short-Term Memory (LSTM), Gate Recurrent Unit (GRU), Bidirectional Long Short-Term Memory (BiLSTM), and Bidirectional Gated Recurrent Unit (BiGRU), were widely employed and achieved notable success in improving the accuracy of NILM. Additionally, this study considered incorporating the attention mechanism to enhance the accuracy of RNN models applied to NILM. Although not directly related to deep learning, the feature integration theory provides a fundamental framework for understanding attention in the human visual system.

Further investigation revealed that contemporary NILM algorithms for air conditioning systems predominantly employ techniques such as MLP and SVM. In contrast, algorithms for other appliances have progressed to incorporate methods like RNNs and attention mechanisms. This disparity underscores the relative dearth of research on NILM for air conditioning systems. Moreover, due to the high sensitivity of air conditioning loads to temperature and humidity, their dynamic adjustments based on indoor and outdoor environmental variations result in significant fluctuations, presenting substantial monitoring challenges. Additionally, current research in the air conditioning NILM field primarily relies on datasets such as UK-DALE and REDD, which need more advanced model development and validation based on real-world data. Consequently, the development of specialized datasets and robust computational models for this field has become a pressing priority.

Therefore, the main focus of this study is as follows:

1)This study constructed the CN-AC-NILM datasets, featuring air conditioning data of China. The dataset was designed to extensively cover electrical data from diverse users and multiple time periods, and also incorporated additional relevant information, such as meteorological parameters and user behavior.

2)The selection of optimal input features for predicting air conditioning load disaggregation models was undertaken through forward feature selection in this research. Various RNN models were then evaluated to compare their performance in NILM applications for air conditioning systems.

3)Attention mechanisms were introduced into the baseline RNN models. Significant enhancements in model accuracy were observed, with LSTM achieving higher accuracy than GRU, and bidirectional models surpassing unidirectional ones in NILM for air conditioning systems.

4)Model accuracy and interpretability were enhanced through the integration of attention mechanisms in this study.

2. MATERIAL AND METHODS

This study aimed to assess the impact of attention mechanisms on RNN model performance for NILM specific to air conditioning systems using the real-world datasets and to develop an optimal RNN-attention models based on the assessment results. The technical roadmap of this study is outlined as follows:

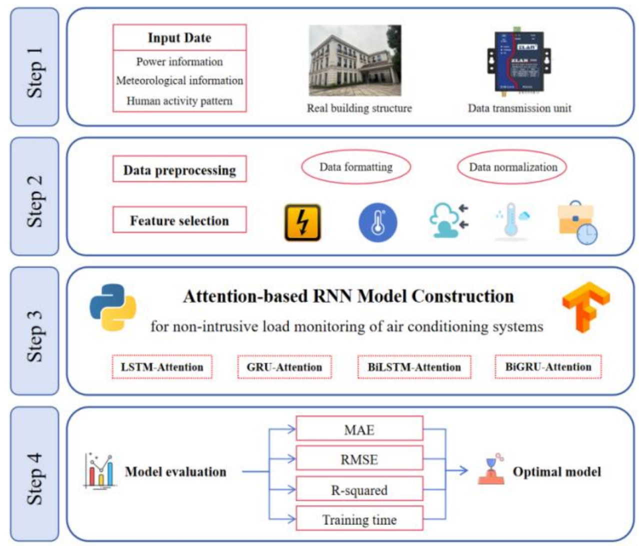

Fig. 1 General framework of this study

First, real-world datasets, which named CN-AC-NILM datasets was obtained as input data by installing monitoring devices on actual building structures, and the input data was preprocessed to improve the quality of the training model. After that, this study employed the forward feature selection method to investigate feature selection for air conditioning load disaggregation and prediction models. Subsequently, four RNN models, including LSTM, GRU, BiLSTM, and BiGRU, were developed for air conditioning load disaggregation and prediction. An evaluation methodology incorporating model accuracy and training time was established to perform a comparative analysis of the comprehensive performance of these four models. Furthermore, based on the constructed LSTM, GRU, BiLSTM, and BiGRU models, attention mechanisms were introduced to develop attention-enhanced variants: LSTM-Attention, GRU-Attention, BiLSTM-Attention, and BiGRU-Attention models. Finally, the performance of the developed models was compared using four evaluation metrics, and the optimal model for NILM specific to air conditioning systems was identified.

2.1 Development of the real-world datasets

To address the shortage of NILM datasets in China, the lack of air conditioning data, and the limited diversity of data features, this study constructed the real-world datasets of China, which named CN-AC-NILM datasets. The datasets are designed to comprehensively cover electrical data from various users and multiple time periods, while also including other crucial information related to electricity usage, such as meteorological parameters and user behavior.

The building selected for this study was a small office building consisting of three floors above ground, each equipped with advanced living facilities and spacious areas. The first and second floors served as the primary areas of the building, both utilizing a multi-split air conditioning system. The first floor was equipped with one GMV-450W/A and one GMV-615W/A multi-split air conditioning unit, while the second floor featured two GMV-400W/A units and one GMV-450W/A unit.

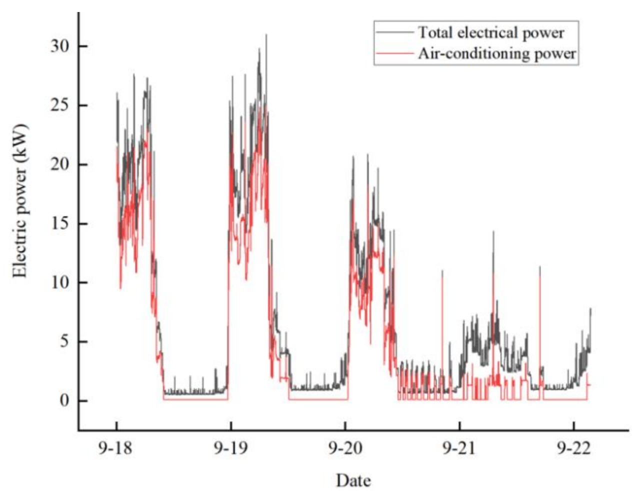

Fig. 2 Hourly electricity consumption data

To construct the CN-AC-NILM dataset, this study installed the corresponding data collection equipment on the first and second floors. These devices were capable of real-time monitoring and recording of electrical load data, meteorological data, and other relevant information. Fig. 2 clearly displays the total residential electrical power and the air conditioning system's electrical power for User 1, located on the first floor, during a specific time period. The results showed that the electrical power of the air conditioning system constituted a significant portion of the total electrical power, highlighting its importance in residential energy consumption.

2.2 Feature selection

This study employed the forward feature selection method for feature selection, considering the high accuracy requirements of the task, the generalizability of the method, the small size of the dataset, and the limited number of features. In the practical experimental procedures, each feature was first trained individually, with the best-performing model saved using a callback function. Specifically, as illustrated in Fig. 3.

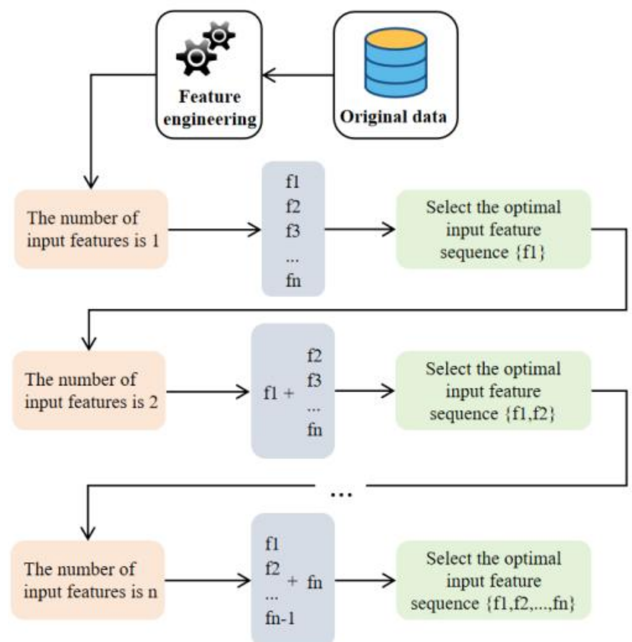

Fig. 3 Flowchart of forward feature selection method

2.3 RNN-Attention Training and Prediction Process

The training data for the RNN-Attention model were inputted based on the previous feature selection conclusions and data preprocessing results. The model training process is shown in Fig. 4. The model training process involved several critical steps to ensure optimal performance. The error calculation phase involved computing the error by comparing the model's predicted output with the proper labels, with a loss function, typically mean squared error, used to measure this discrepancy. During the training process, the model's performance on the validation set was evaluated at each epoch; if the validation set error was smaller than that of the previous model, the model was saved, and training continued. Additionally, the training process monitored whether the number of iterations had reached the predefined limit, set to 100 in this study's experiments. If the maximum number of iterations were reached, the training process would terminate. Following error calculation, the backpropagation algorithm was used to update the model's parameters to minimize the loss function. Backpropagation is an effective optimization method that adjusts model parameters along the gradient direction of the loss function, enabling the model to progress toward a more optimized state.

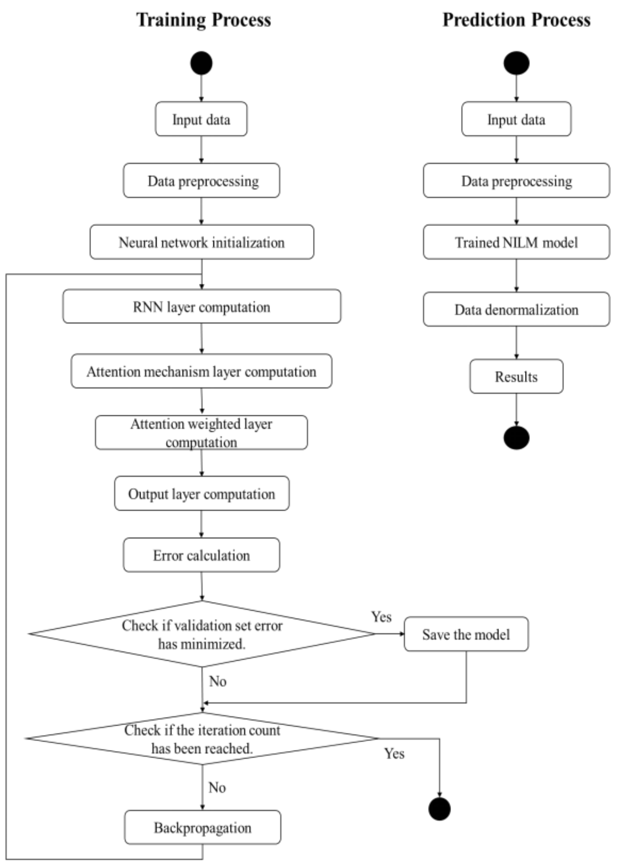

Fig. 4 RNN-Attention training and prediction process

2.4 Model Implementation

In the experiments, hyperparameters for all models were standardized to ensure the comparability of results. Expressly, the hidden neuron count for LSTM, GRU, BiLSTM, and BiGRU layers was set to 64. The number of iterations was fixed at 100, the loss function employed was Mean Squared Error (MSE), the optimizer used was Adam, the batch size was set to 64, the time step was 10, and the learning rate was set to 0.001.

All experiments were conducted on a device equipped with an NVIDIA GeForce RTX 3060 Laptop GPU. Python 3.9 and TensorFlow 2.10 were utilized as the deep learning framework, and cuDNN version 8302 was employed to accelerate the model training process.

3. RESULTS AND DISCUSSION

3.1 Reak-world Data Processing

Real-world data from two office users in China, were selected in this study. The data spans from September 18, 2023, 9:55 AM, to September 22, 2023, 11:00 AM, with a sampling interval of 5 seconds. The data underwent a preprocessing procedure that included data cleansing, data formatting, and data normalization. Each record contained eight independent variables: total power consumption, outdoor dry-bulb temperature, outdoor relative humidity, outdoor wet-bulb temperature, outdoor wind speed, outdoor wind direction, rainfall, and working time coefficient, along with one dependent variable: air conditioner power consumption. The results of the feature selection experiments are shown in Fig. 5.

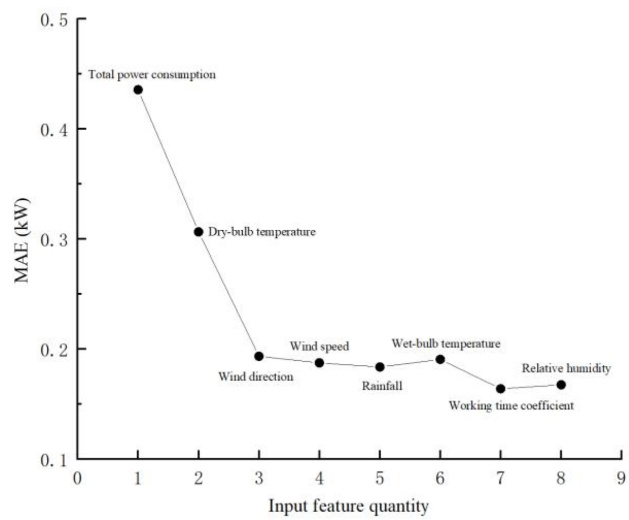

Fig. 5 MAE of the best performing model for each number of input features

3.2 *Baseline Model Comparison*

The performance of four baseline models—LSTM, GRU, BiLSTM, and BiGRU — was examined. Fig. 6 displays the air conditioning load disaggregation results of these benchmark models.

BiLSTM achieved the highest accuracy, while GRU exhibits the lowest. In terms of training time, GRU had the shortest training time, and BiLSTM had the longest. Further analysis revealed that LSTM-based models surpassed GRU-based ones in accuracy. Further analysis revealed that LSTM was the optimal choice due to its excellent balance between accuracy and training time. Additionally, bidirectional structure models outperformed unidirectional ones, with reductions in MAE of 1.16% for LSTM (BiLSTM vs. LSTM) and 2.89% for GRU (BiGRU vs. GRU). Notably, the improvement with bidirectional structures was limited, possibly because air conditioning NILM tasks relied more on forward information, and the contribution of reverse information was limited.

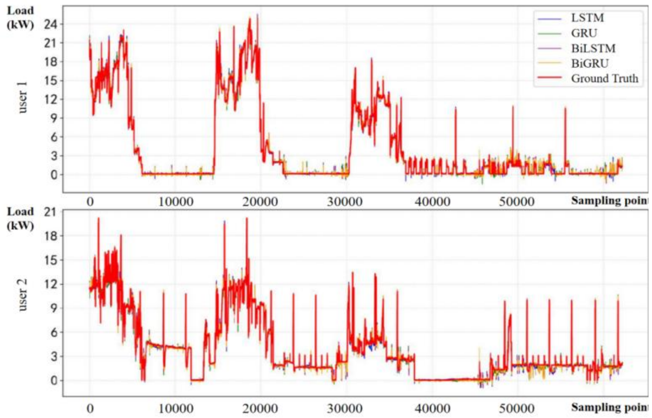

Fig. 6 Decomposition effect of air conditioning load of each RNN model

3.3 Model evaluation with the introduction of attention mechanism

The impact of introducing the attention mechanism on the performance of RNN models is explored. Fig. 7 display the partial air conditioning load decomposition results for RNN models and their RNN-Attention counterparts, revealing the superior decomposition performance of the Attention model.

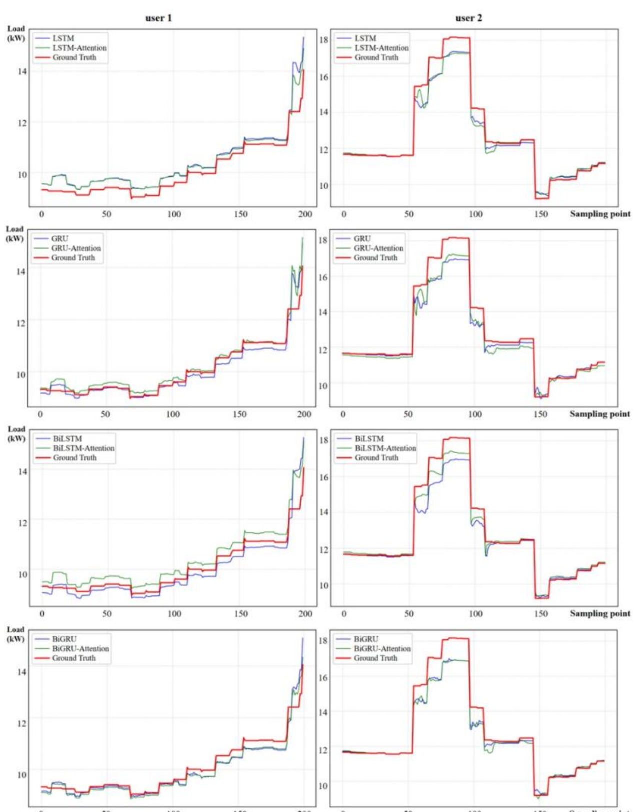

Fig. 7 The partial air conditioning load decomposition results of RNN-Attention models

To investigate the performance variations of models incorporating Attention mechanisms using different base units, LSTM and GRU were compared, as well as BiLSTM and BiGRU. Results indicate that, in terms of both accuracy and training time, the LSTM-based model outperforms the GRU-based model after introducing the Attention mechanism.

3.4 Interpretability Analysis

Enhancing the interpretability of deep learning models was crucial for assessing their performance. Unlike traditional benchmark models, which were often opaque, attention models utilizing attention mechanisms offered a promising solution. This study analyzed attention weights from various models to understand how they prioritized information when decomposing air conditioning loads. By examining these weights, insights were gained into how models processed time series data and emphasized specific segments of the input sequence. This analysis clarified how models utilized input information to make predictions, thereby shedding light on their behavior.

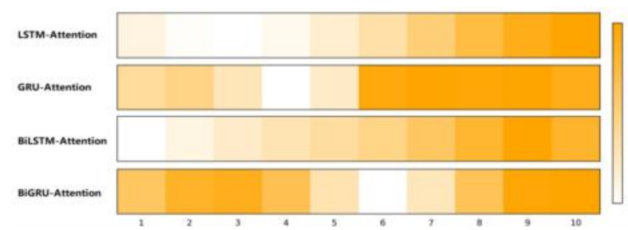

Fig. 8 Heatmap of attention weights in models

The model inputs data from the first 10 time steps (including the current one) in a (10, 7) format. Fig. 8 depicts the attention weight distributions of four models during the air conditioning NILM task. The x-axis represents time steps, and the y-axis represents different Attention-based models, with color intensity reflecting attention weight magnitude (darker colors indicating higher weights).

4. CONCLUSIONS

This study established RNN-attention models to improve the accuracy of NILM for air conditioning systems. These models include LSTM, GRU, BiLSTM, BiGRU, LSTM-Attention, GRU-Attention, BiLSTM-Attention, and BiGRU-Attention, which were trained and evaluated on real-world data, using MAE, RMSE, R2, and training time as evaluation metrics. The main conclusions are as follows:

1)This study constructed the CN-AC-NILM datasets, featuring air conditioning data of China. This dataset not only provides practical cases for energy management but also contributes valuable data resources and research outcomes to the development and application of NILM technology.

2)This study utilized forward feature selection to optimize the input features for the air conditioning load decomposition prediction model. The findings suggest that incorporating total power consumption, outdoor dry-bulb temperature, wind direction, wind speed, rainfall, outdoor wet-bulb temperature, and working time coefficient significantly enhances prediction accuracy. This demonstrates the importance of including both electrical and environmental factors in load monitoring for better energy management.

3)Introducing the attention mechanism into RNN models not only improves performance but also enhances model interpretability. Experimental results show that this approach boosts model accuracy by an average of 21.09%, with only a modest 5.33% increase in training time. This trade-off is practical and offers substantial real-world benefits, such as more accurate load disaggregation, leading to better energy usage insights and targeted energy-saving strategies, which can reduce operational costs for both residential and commercial buildings.

4)In terms of neural network units, LSTM models outperform GRU models in accuracy, although they require longer training times. When combined with the attention mechanism, LSTM-Attention models achieve the best overall performance. Additionally, bidirectional structures (BiLSTM, BiGRU) outperform their unidirectional counterparts, though at the cost of increased training time. By enhancing accuracy in load monitoring, these models can help in optimizing HVAC system performance, reducing energy waste, and lowering electricity bills, which translates to significant financial and environmental gains in large-scale deployments.

Although the proposed attention-enhanced RNN models achieved significant improvements in air-conditioning NILM accuracy and interpretability, several limitations remain. The current CN-AC-NILM dataset covers only five days of monitoring data from a small office building, which limits seasonal and typological diversity. Nevertheless, the high-resolution 5-second records effectively capture typical operating patterns. Future work will expand the dataset across different seasons and building types and evaluate model performance under lower sampling frequencies (e.g., 1-minute) to enhance generalizability and practical applicability.

In addition, further interpretation of the attention mechanism revealed that higher attention weights corresponded to periods of startup or load fluctuation, indicating that the model effectively focuses on transient behaviors critical for load disaggregation. Future studies will strengthen this interpretability analysis by linking attention patterns with specific air conditioning operations, thereby improving the physical interpretability and deployment potential of NILM models in real-world energy management.

This work was supported by the National Key R&D Program from the Ministry of Science and Technology, China [grant number 2022YFC38015031].

REFERENCE

[1] J. Ngarambe, G.Y. Yun, M. Santamouris, The use of artificial intelligence (AI) methods in the prediction of thermal comfort in buildings: Energy implications of AI-based thermal comfort controls, Energy and Buildings 211 (2020) 109807.

[2] Y. Himeur, A. Alsalemi, F. Bensaali, A. Amira, Efficient multi-descriptor fusion for non-intrusive appliance recognition, in: 2020 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, 2020: pp. 1 – 5.

[3] W.W. Che, C.Y. Tso, L. Sun, D.Y. Ip, H. Lee, C.Y. Chao, A.K. Lau, Energy consumption, indoor thermal comfort and air quality in a commercial office with retrofitted heat, ventilation and air conditioning (HVAC) system, Energy and Buildings 201 (2019) 202 – 215.

[4] Y. Himeur, K. Ghanem, A. Alsalemi, F. Bensaali, A. Amira, Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives, Applied Energy 287 (2021) 116601.

[5] S. Su, Y. Yan, H. Lu, L. Kangping, S. Yujing, W. Fei, L. Liming, R. Hui, Non-intrusive load monitoring of air conditioning using low-resolution smart meter data, in: 2016 IEEE International Conference on Power System Technology (POWERCON), IEEE, 2016: pp. 1 – 5.

[6] Y. Kim, S. Kong, R. Ko, S.-K. Joo, Electrical event identification technique for monitoring home appliance load using load signatures, in: 2014 IEEE International Conference on Consumer Electronics (ICCE), IEEE, 2014: pp. 296 – 297.

[7] M. Aiad, P.H. Lee, Unsupervised approach for load disaggregation with devices interactions, Energy & Buildings 116 (2016) 96 – 103.

[8] G.W. Hart, Nonintrusive appliance load data acquisition method: Progress report, MIT Energy Laboratory, 1984.