

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/168352/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Zhang, Chengyu, Luo, Zhiwen , Rezgui, Yacine and Zhao, Tianyi 2024. Enhancing multi-scenario data-driven energy consumption prediction in campus buildings by selecting appropriate inputs and improving algorithms with attention mechanisms. *Energy and Buildings* 311 , 114133. 10.1016/j.enbuild.2024.114133

Publishers page: <http://dx.doi.org/10.1016/j.enbuild.2024.114133>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Enhancing multi-scenario data-driven energy consumption prediction in campus buildings by selecting appropriate inputs and improving algorithms with attention mechanisms

Chengyu Zhang¹, Zhiwen Luo², Yacine Rezgui³, Tianyi Zhao^{1*}

1 Institute of Building Energy, Dalian University of Technology, Dalian, China

2 Welsh School of Architecture, Cardiff University, Cardiff, United Kingdom

3 School of Engineering, Cardiff University, Cardiff, United Kingdom

* Corresponding author: zhaotianyi@dlut.edu.cn

Abstract: Effective building energy prediction is vital for sustainable development, especially with an increasing focus on flexibility and elasticity in building energy usage. However, challenges persist in input and algorithm selection including suitable input selection and the balance of calculation time consumption and prediction precision. This study proposed a method for suitable input selection and introduced a CNN-LSTM-SE model for enhancing prediction precision. The CNN-LSTM-SE model integrates a convolutional neural network (CNN) process and a squeeze-and-excitation (SE) block attention mechanism into a long short-term memory (LSTM) neural network. Experiments across seven buildings demonstrate that selecting inputs improves the prediction performance with a 13.94% decrease in MAPE (mean absolute percent error), a 3.36% increase in R^2 (coefficient of determination), and an 11.20% decrease in CV-RMSE (coefficient-of-variation of root mean square error) on average compared to not selecting inputs. Moreover, implementing CNN-LSTM-SE yields a 21.44% MAPE decrease, a 6.25% R^2 increase, and a 21.82% CV-RMSE decrease compared to alternative algorithms. Simultaneously employing suitable inputs and CNN-LSTM-SE balances time consumption and prediction precision optimally.

Keywords: Building energy consumption prediction; Building energy branch; Input selection; SE block attention mechanisms

Abbreviations list

ANN	Artificial neural network
CNN	Convolutional neural network
SVM	Support vector machine
ARIMA	Autoregressive integrated moving average
DT	Decision tree
GA	Genetic algorithm
LSTM	Long short-term memory (neural network)
SE	Squeeze-and-excitation (attention mechanisms)
r	Correlation coefficient
R ²	The coefficient of determination
MAPE	Mean absolute percentage error
CV-RMSE	Coefficient-of-variation of root mean square error
PSO	Particle swarm optimization
AC	Air conditioning
CV	Coefficient of variation
NCV	The coefficient of variation for the data set with the normalization process
IQR	Interquartile range
NIQR	The interquartile range for data set with the normalization process
M	Proportion coefficient to measure the proportion of each energy consumption branch or item
AM	Average of M with an hourly time-scale
STDEVP	The standard deviation of a data set
MEAN	The mean value of a data set
Max	The maximum value in a data set
Min	The minimum value in a data set

1 Introduction

The global energy crisis and increasing environmental concerns have prompted the need for sustainable cities and societies. Buildings, responsible for over one-third of total energy consumption across the globe ^[1], account for more than 40% of energy consumption in the US and EU ^[2]. Similarly, in China, building energy consumption constitutes 45.5% of national energy consumption, with material production, construction, and operation phases contributing 22.3%, 1.9%, and 21.3% respectively ^[3]. These figures highlight the urgency of managing building energy consumption and reducing carbon emissions. To address this challenge, effective and real-time energy management technologies are vital, requiring an online accurate energy consumption prediction system with high accuracy. Various approaches such as the promotion of distributed energy systems ^[4], carbon trading mechanisms ^[5], and smart city construction ^[6] emphasize the significance of establishing accurate and online energy consumption predictions. Furthermore, the introduction of emerging technologies like distributed energy systems complicates energy supply-demand matching, making precise and immediate energy consumption prediction at an hourly scale critical for energy network scheduling. Moreover, some individuals and households ^[7] can shift the energy consumption of some appliances to off-peak periods, such as washing machines, which can reduce energy consumption and lower electricity costs ^[8–10].

Especially in recent years, there has been a growing emphasis on addressing the issues of flexibility and elasticity in building energy usage, extending beyond the realm of building energy consumption prediction alone. Simultaneous predictions of whole building energy consumption, as well as energy consumption prediction in air conditioning systems, lighting systems, and other electrical devices connected to wall outlets, have become increasingly important. These predictions hold immense significance as they enable better control over or adjust human behaviors in utilizing these devices. It can help to shift the energy consumption of some appliances to off-peak periods. Furthermore, it can align energy demand with variable energy supply, such as solar and wind power which are subject to temporal and meteorological fluctuations. And then it will be helpful for the flexibility of building energy uses. Several research studies have explored the importance of energy consumption prediction and its impact on energy management and demand response strategies. For instance, Zhan et al. improved energy flexibility and PV self-consumption for a tropical net zero energy office building based on four kinds of forecasting for outdoor temperature, solar irradiance, PV generation, and plug loads ^[24]. Khalil et al. comprehensively reviewed the relevant studies and proposed that forecasting building energy consumption can support energy demand side management, balance energy supply and demand, and enable the energy flexibility of buildings to support the energy network ^[25]. Therefore, by incorporating accurate energy consumption prediction models with different items (including ACs, lights, etc.) into building management systems,

stakeholders can make informed decisions, enhance energy efficiency, and contribute to a more sustainable built environment.

The field of building energy consumption prediction comprises three primary approaches: physical modeling methods, combinatorial methods, and data-driven methods. Physical modeling methods rely on simulation software which incorporates climate files, building geometry, occupant activity schedules, etc. However, limitations in reflecting real-time interactions among occupants, energy systems, and the environment hinder its effectiveness, especially in some online applications. Moreover, questions have been raised regarding their accuracy ^[11–12]. Combinatorial methods integrate physical modeling and data-driven approaches, leading to improved performance. However, using the software can also compromise online prediction performance, and it also has some other challenges such as a lack of theoretical hypotheses and standardized nomenclature ^[14–15]. Considering the aims, this study will focus on data-driven methods.

As for data-driven methods, it relies on sufficient data and its main work may lie in developing and selecting suitable inputs and effective algorithms. This study attempted to summarize and analyze some recent related studies as follows.

As for input system selection, the possible parameters mainly include indoor environment-related parameters (indoor air temperature ^[38], indoor air relative humidity ^[39], and room carbon dioxide concentration ^[38], etc.); outdoor meteorological parameters (outdoor air temperature ^[38], outdoor relative humidity ^[39], horizontal solar radiation ^[38], wind speed ^[40], comprehensive air-quality index ^[39], rain ^[38], etc.); occupant-related parameters (occupancy ^[48], number of occupants ^[49], stay hours ^[50], Time interval of energy usage activities ^[51], occupant density ^[38], age of the users ^[41], gender of the users ^[43], indoor thermal comfort ^[44], and cultural background ^[40], etc.); equipment-related parameters (type of ACs ^[41], heating and cooling set point ^[50], equipment density ^[47], equipment schedules ^[52], operation duration of ACs ^[42], etc.); time-related parameters (time of the day ^[38], and other events ^[40], etc.); and building information parameters (whether the house is old or new ^[41], building geometry ^[45], shape ^[46], orientation ^[46], roof type ^[46], etc.).

However, there are still important challenges to the selection of current inputs. First, different energy consumption predictions may require different inputs. For example, theoretical analysis shows that lighting energy consumption has a strong correlation with indoor illuminance but a weak correlation with indoor air temperature, while air conditioning energy consumption shows the opposite. Another example may be that plug-load (including lights, sockets, and some split ACs connected to wall outlets) energy consumption may be strongly influenced by time while centralized ACs may not. Consequently, there is a pressing requirement to propose methods and conclusions for selecting suitable inputs when forecasting energy consumption in different cases including building, AC, lighting, plug loads, etc. Second, even in the same scenario, the selection of inputs

often varies. Moreover, although some recent studies discussed the suitable parameters affecting building energy consumption ^[36–37], many different types of inputs are not considered in some studies. For instance, despite forecasting energy consumption for office buildings ^[53–56], some studies do not consider indoor air temperature, while others regard it as a significant influencing factor. Furthermore, some studies consider various time elements (such as disparities of hours, months, etc.), while others scarcely consider them. It must be noted that these variances do not necessarily denote the incorrectness of these studies, and such differences may arise from distinct building characteristics. These characteristics not only refer to the building type but also encompass energy consumption compositions (such as the proportion of lighting energy consumption to total energy consumption), energy consumption features (such as the intensity of fluctuations), and the climate zone in which the building is located.

For these reasons, it is almost impossible to draw a universal consensus conclusion, as it is difficult to have a study that encompasses all types of architecture from around the world. However, it is possible to propose a simple, universal method of input optimization as a standard step in energy consumption prediction. In this way, the researcher can quickly select appropriate inputs when oriented to different predictions. This is not only for energy consumption prediction but also to help analyze the energy-saving potential of different systems, to calculate the energy-saving rate after certain strategies, etc.

As for the algorithms and their optimizations of data-driven building energy consumption prediction, the feasibility of many algorithms has been verified. For example, back propagation neural networks ^[57], recurrent neural networks ^[59], convolutional neural networks ^[58], long short-term memory neural networks ^[59], support vector machines ^[60], multiple linear regression ^[60], autoregressive integration and moving average ^[61], decision trees ^[63], genetic algorithms ^[62], etc. Referring to the statistics of related studies ^[13], when using intelligence algorithms, 47%, 25% and 4% of the studies used ANN, SVM, and DT respectively to train the models. While the remaining 24% of the studies used statistical algorithms. However, each of these basic algorithms has some challenges, such as the tendency to fall into local optimum, long calculation time, difficulty in convergence, low prediction accuracy, etc. Considering their applicability and performance characteristics in energy consumption prediction, consistent with existing studies, this study tries to use long short-term memory (LSTM) as the basic algorithm in the face of such time-series problems as hourly energy consumption prediction.

To further improve the accuracy, speed up the convergence, avoid overfitting, etc., many studies additionally use many kinds of optimizations on top of the basic algorithm. First, the data processing process can be optimized with optimization algorithms, such as using K-means clustering methods ^[64] to classify the input and then predict the energy consumption by different

scenes. Some other methods such as using generative adversarial networks (GAN) to obtain generating parallel data, and then training the prediction model using a hybrid dataset including both original and generating parallel data ^[66]; using methods such as Minimum redundancy maximum relevance (MRMR) algorithm to analyze the data features and then perform classification prediction ^[67]. Second, the hyperparameter optimization selection process can be optimized algorithmically, for example, using some swarm intelligence algorithms for hyper-parameters tuning such as weights and thresholds of neural networks ^[29–35], including sparrow search algorithm ^[68], particle swarm optimization ^[69], cuckoo search algorithm ^[70], ant colony optimization ^[71], grey wolf optimization ^[72], and whale optimization algorithm ^[73], etc. Finally, it can also be optimized with a combination of multiple algorithms. For example, Liang et. al. combined the deep ensemble (DE) model and autoregressive (AR) model ^[65], and it is worth noting that the two previous types of optimization processes can also be used in combination at the same time.

However, there are still important challenges to algorithms. The key challenge is to balance the calculation time consumption and prediction precision, expanding to include: the tendency to fall into local optimum, the complexity or long-time computation of the procedure, the low prediction accuracy, the difficulty of online application, the difficulty of reflecting both peak and trough values, and the difficulty of applying to various types of buildings, etc. Specifically, for example, some swarm intelligence algorithms may require a lot of extra time for iteration which makes online applications difficult; some algorithms may not be suitable for buildings with complex energy consumption fluctuations; and some algorithms may stabilize the predicted values around the mean value but have weak predictions of peaks and valleys, etc.

For these reasons, it is necessary to propose a training algorithm structure that brings a large improvement in prediction accuracy with as little a consumption of extra time as possible (compared to an unoptimized LSTM).

To address the above challenges on inputs and algorithms, this study proposes a complete building energy consumption prediction system mainly including selecting suitable inputs and upgrading algorithms. First, this study used correlation coefficient analysis to discuss the most appropriate inputs for energy consumption prediction of the whole building, air conditioning, and plug loads. Second, this study introduces the CNN-LSTM-SE algorithm, which integrates CNN into LSTM for optimizing the data processing process, and introduces the Squeeze-and-Excitation block attention mechanism (SE) into LSTM for optimizing the hyperparameter optimization process. Finally, in this study, eighty-four sets of data simulation experiments consisting of seven different buildings, two inputs, and six algorithms are cross-composed for validation, and the changes in prediction performance due to improved inputs and improved algorithms are discussed. The two inputs include suitable parameters with selection by the proposed method and all parameters. The

six algorithms include: the most conventional LSTM, the proposed CNN-LSTM-SE, CNN-LSTM, and LSTM-SE used to validate that the combination of the CNN and SE modules is valuable, as well as the optimized algorithms LSTM-PSO and LSTM-GA, which are commonly used nowadays.

The remaining sections are organized as follows. Section 2 provides all the methods and necessary information for this study. Section 3 presents the results on selecting suitable inputs and forecasting energy consumption in different scenes. Section 4 will discuss the performance changes when enhancing the energy consumption prediction, and list the limits of this study. In Section 5, the conclusions are presented. The targets and overall research methods are shown in Figure 1.

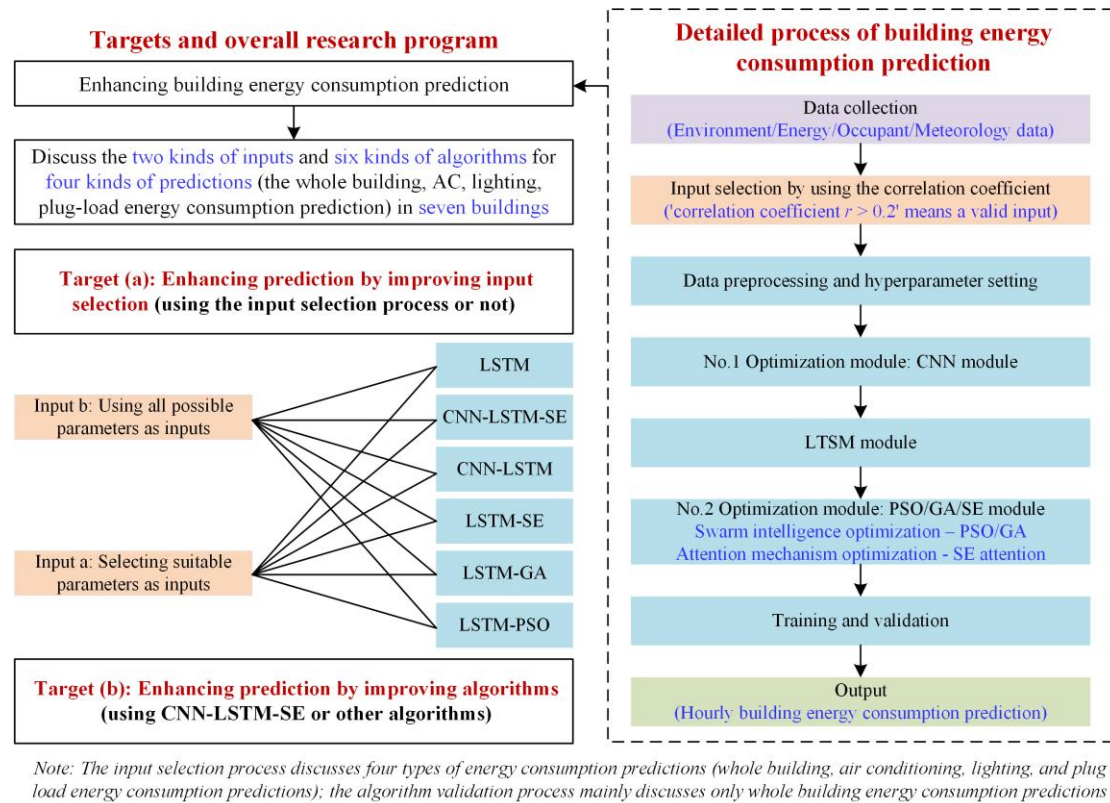


Figure 1 The overall targets, methods, and research flowchart of this study

2 Methodology

2.1 Sample buildings and related data collection

The sample buildings in this study are shown in Table 1, coded Buildings A–G. The main types of campus buildings include dormitories, office buildings, educational buildings, libraries, and academic laboratory buildings. The energy consumption pattern of dormitory buildings is relatively homogeneous, and the dormitories in the target campus do not have air-conditioning equipment and have restricted electricity consumption time (power off after 22:30). Therefore, the dormitory buildings of this campus do not seem to be typical. Moreover, since there are more studies on dormitory or residential buildings, this study chooses some typical office buildings, educational buildings, libraries, and academic laboratory buildings as study cases (Buildings A–G). These

buildings will have an EUI (Electrical consumption intensity) of more than 30 kWh/m², and these buildings will have as much energy-using equipment as possible and can be freely used.

Table 1 The sample buildings in this study

Codes	Types	Energy consumption items	Heating and cooling systems	EUI (kWh/m ²)
Building A	Laboratory	Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water and using split ACs as a supplement Cooling: split ACs	75.76 kWh/m ²
		Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water and using split ACs as a supplement Cooling: variable refrigerant flow multilink ACs	
Building B	Office	Centralized ACs, and Plug-load (including lights, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water alone Cooling: centralized ACs (with screw chillers)	44.89 kWh/ m ²
		Centralized ACs, and Plug-load (including lights, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water alone Cooling: centralized ACs (with screw chillers)	
Building C	Laboratory	Centralized ACs, and Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water and using some split ACs as a supplement Cooling: centralized ACs (with screw chillers)	41.55 kWh/ m ²
		Centralized ACs, and Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water and using some split ACs as a supplement Cooling: centralized ACs (with screw chillers)	
Building D	Library	Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water alone Cooling: none in most rooms and few split ACs in computer and discussion classrooms (only three)	30.54 kWh/ m ²
		Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water alone Cooling: none in most rooms and few split ACs in computer and discussion classrooms (only three)	
Building E	Library	Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water alone Cooling: none in most rooms and few split ACs in computer and discussion classrooms (only three)	60.09 kWh/ m ²
		Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water alone Cooling: none in most rooms and few split ACs in computer and discussion classrooms (only three)	
Building F	Education	Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water alone Cooling: none in most rooms and few split ACs in computer and discussion classrooms (only three)	62.74 kWh/m ²
		Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water alone Cooling: none in most rooms and few split ACs in computer and discussion classrooms (only three)	
Building G	Education	Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water alone Cooling: none	37.76 kWh/m ²
		Plug-load (including lights, split ACs, and other electrical equipment connected to wall outlets)	Heating: District heating system with municipal hot water alone Cooling: none	

An energy-consumption monitoring and management system was installed in the sample buildings, which can comprehensively record the current, voltage, and energy consumption changes in all energy-use items and branches and upload related data to cloud platforms, as shown in Figure 2(a). Indoor environmental data were obtained from ceiling-mounted sensors installed in the buildings with at least one sensor in each room and corridor, where the installation location and equipment schematic are shown in Figure 2(b). Outdoor meteorological data were recorded and collected from the China Meteorological Science Data Center ^[16]. Occupancy and the number of occupants were obtained from the automatic access control system or similar monitoring systems (installed in Buildings A, B, C, F, and G) and the seat reservation system (installed in Buildings D

and E) in 2021 and collected hourly, as shown in Figure 2(c). Moreover, this study also used indoor CO₂ concentrations reflecting the occupancy or number of occupants in some buildings, and this alternative has also been proven effective in many studies [17–21].

All data were collected with an hour-scale in 2021. In universities, various activities of students and teachers are influenced by campus schedules, such as course schedules and examination plans. As for a year, it generally includes vacation, exam days, and non-exam days. As for the day of a period, it can be divided into two kinds including weekdays and weekends. As for the time of one day, it generally includes working time, transition time, and night.

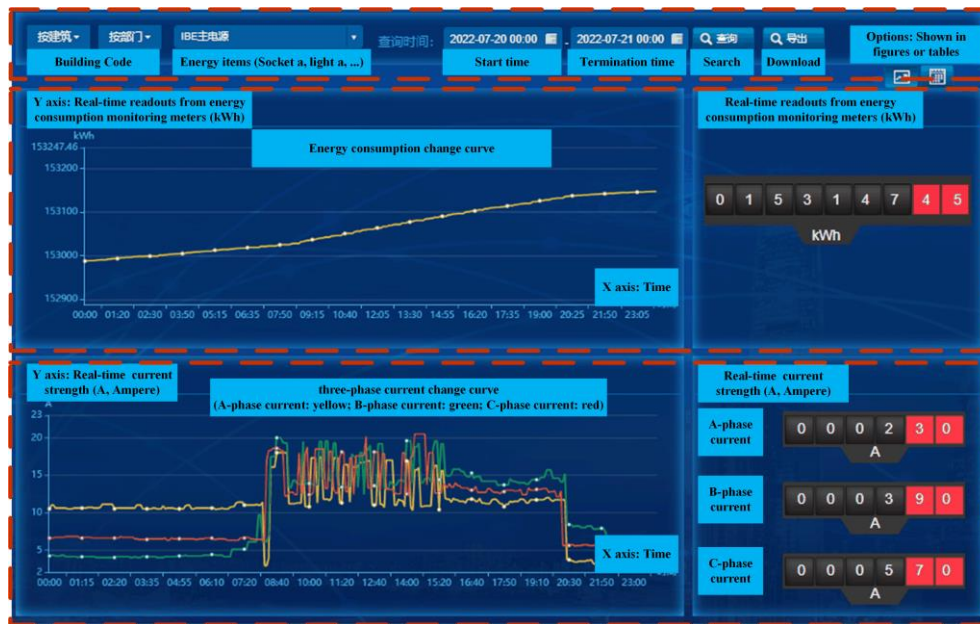


Figure 2(a) The monitoring and management of building energy consumption [74]

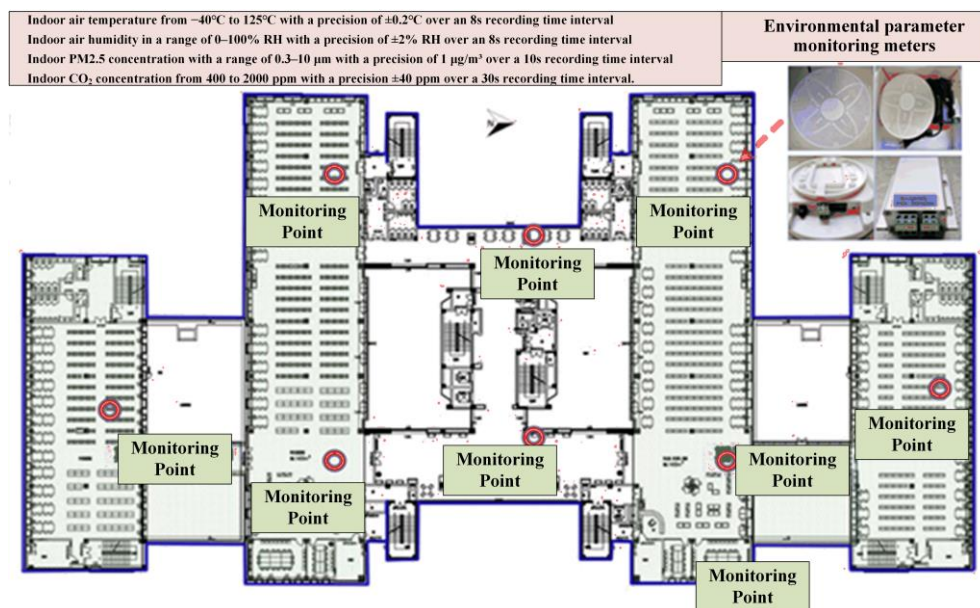


Figure 2(b) The monitoring and management of building indoor environmental parameters [74]

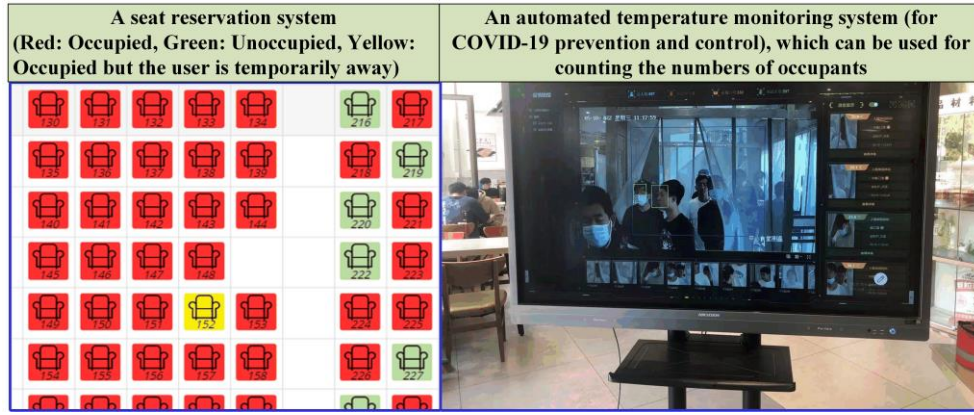


Figure 2(c) The monitoring and management of occupancy (or number of occupants) ^[74]

Figure 2 Different parameter monitoring systems in this study (Several previous studies have applied similar data collection platforms ^[74])

2.2 Characteristics of energy consumption for different items

To measure the fluctuation complexity of building energy consumption and the proportion of each branch in the total energy consumption, this study adopts the coefficient of variation (CV) and interquartile range (IQR) to assess energy consumption variability and changes. The proportion coefficient (M) is also defined to measure the proportion of each energy consumption branch or item. To ensure a more objective comparison among different buildings, this study normalizes the building energy consumption using the Max-Min normalization method before analyzing the proportions and complexities. The parameters obtained from normalization are defined as NCV , $NIQR$, and NM . The proportion coefficient NM is further taken as the average value to reflect the average proportion of each energy consumption branch, defined as ANM . The equations are shown as follows:

$$CV ([x_1, x_2, x_3 \dots, x_n]) = \frac{STDEVP([E_{x1}, E_{x2}, E_{x3} \dots, E_{xn}])}{MEAN ([E_{x1}, E_{x2}, E_{x3} \dots, E_{xn}])} \quad (1)$$

$$IQR ([x_1, x_2, x_3 \dots, x_n]) = Q_3 ([x_1, x_2, x_3 \dots, x_n]) - Q_1 ([x_1, x_2, x_3 \dots, x_n]) \quad (2)$$

$$M_t = \frac{E_{x,t}}{E_{total,t}} \times 100\% \quad (3)$$

$$N_p = Normalization ([x_1, x_2, x_3 \dots, x_n]) = \frac{x_p - Min([x_1, x_2, x_3 \dots, x_n])}{Max([x_1, x_2, x_3 \dots, x_n]) - Min([x_1, x_2, x_3 \dots, x_n])} \quad (4)$$

$$NCV = CV(Normalization ([x_1, x_2, x_3 \dots, x_n])) \quad (5)$$

$$NIQR = IQR (Normalization ([x_1, x_2, x_3 \dots, x_n])) \quad (6)$$

$$AM = \frac{\sum_{t=1}^N M_t}{N} \quad (7)$$

Where $[x_1, x_2, x_3 \dots, x_n]$ represent a data set, where x_p represents an individual data point in the column. The definitions for other data sets (CV, IQR, M, and E) are the same. E_x and E_{total} represent the energy consumption of an item and the total energy consumption respectively. Max and Min represent the maximum and minimum values in a data set. $STDEVP$ and $MEAN$ respectively

represent the standard deviation and the mean value of a data set. Q_3 and Q_1 represent the third and first Quartile respectively. A quartile is a value that is arranged in statistics from the smallest to the largest and divided into four equal parts, at three division points. Q_1 , also named the lower Quartile, is equal to the 25th percentile of the sorted data. Q_3 , also named the upper Quartile, is equal to the 75th percentile of the sorted data. The results on characteristics of energy consumption for different items by NCV , $NIQR$, and NM are shown in Table 2.

Table 2 The results on characteristics of energy consumption for different items by NCV , $NIQR$, and AM

	NCV_t	NCV_p	NCV_l	NCV_a	$NIQR_t$	$NIQR_p$	$NIQR_l$	$NIQR_a$	AM_p	AM_l	AM_a
A	0.244	0.168	0.278	0.097	0.291	0.196	0.293	0.719	57.23%	38.44%	4.33%
B	0.203	0.131	0.210	0.156	0.823	0.706	0.424	1.122	3.68%	64.74%	31.58%
C	0.508	0.164	0.250	0.725	0.506	0.203	0.234	1.036	22.96%	50.99%	26.05%
D	0.315	0.519	0.301	0.009	0.719	0.536	0.981	2.602	69.02%	24.75%	6.23%
E	0.317	0.543	0.325	0.241	0.958	1.011	0.880	1.032	9.64%	70.18%	20.18%
F	0.240	-	-	-	0.333	-	-	-	100%	-	-
G	0.658	-	-	-	0.875	-	-	-	100%	-	-

Note: The subscripts t , p , l , and a represent the parameters used to analyze the whole building, plug-load, light, and AC energy consumption data sets, respectively.

2.3 Correlation coefficient analyses

To determine reasonable inputs, the main factors affecting energy consumption are typically used as input parameters. In determining the influencing factors, the correlation coefficient r is commonly used to measure the correlation. Generally, for fitting performance, " $r = 0$ " indicates no correlation, " $0 < |r| \leq 0.3$ " indicates a weak correlation, " $0.3 < |r| \leq 0.5$ " indicates a low correlation, " $0.5 < |r| \leq 0.8$ " indicates a significant correlation, " $0.8 < |r| < 1$ " indicates a high correlation, and " $r = 1$ " indicates perfect linear correlation [22–23]. However, these thresholds are not absolute and should be adjusted according to specific circumstances. In the field of building energy research, building energy consumption is a variable that receives the influence of various parameters such as occupant-related parameters, indoor environmental parameters, outdoor meteorology parameters, equipment-related parameters, time, etc., so it is difficult to require that the correlation between each parameter and the energy consumption is greater than 0.8 (high correlation) [22–23]. Currently, the correlation threshold commonly used in recent research generally is 0.1–0.5 (Specific values include 0.1, 0.2, 0.3, or 0.5 [81–85]) in building electricity, heating, or cooling consumption prediction. Referring to these studies [22–23, 81–85], in this study, if the correlation coefficient is greater than 0.2, it can be considered a potential input parameter. The correlation coefficient (r) is calculated as shown in equation (8):

$$r(a, b) = \frac{Cov(a, b)}{\sqrt{Var(a) \times Var(b)}} \quad (8)$$

Where a and b are different variables, Cov and Var are the covariance and variance process.

2.4 CNN and LSTM

The LSTM model utilizes gating mechanisms to effectively retain historical information, addressing the issue of vanishing gradients in shallow neural networks^[26–27]. CNNs, known for their abstract representation capabilities, are commonly employed for image recognition and have also been adapted for extracting hidden features from input data^[26–27]. Considering that both LSTM and CNN represent classical neural network architectures, this study does not detail the mathematical derivation process but presents only the results of the detailed hyperparameter selection process.

For the CNN hyperparameters^[26–27], the maximum number of iterations (MaxEpochs) is set to 100, with an initial learning rate (InitialLearnRate) of 0.01. Additionally, LearnRateDropPeriod is 80, and LearnRateDropFactor is 0.8 (The learning rate is multiplied by a factor every time a certain number of periods elapse, where the multiplication factor is LearnRateDropFactor and the number of periods is LearnRateDropPeriod). In the case of LSTM^[26–27], the LearnRateDropPeriod is 800, and the LearnRateDropFactor is 0.8. MaxEpochs is set to 1000, and InitialLearnRate is 0.01. L2 Regularization^[28] is implemented in both CNN and LSTM models to prevent overfitting. A mean-squared-error loss function and a stochastic gradient descent (SGD) were used to calculate the loss and optimization. Moreover, the *ReLU* function (Linear rectification function) will be used as the activation function in the CNN process, of which the convolutional kernel size is 64^[79]. The *tanh* function will be used as the activation function in the LSTM process. The neurons of LSTM are set to 32 and the hidden layers are set to 2 (Note: In LSTM-PSO and LSTM-GA, PSO and GA are used to solve for the optimal number of hidden layers instead of setting it to 2 directly).

2.5 Particle swarm optimization (PSO) and genetic algorithm (GA)

Particle Swarm Optimization (PSO) is inspired by social behavior such as bird flocking or fish schooling. In PSO, a population of potential solutions, called particles, move through the solution space following the best particles. Each particle adjusts its position according to its own experience, as well as the experiences of its neighbors. On the other hand, Genetic Algorithms (GA) are based on the principle of natural selection and genetics. This algorithm uses the concepts of selection, crossover, and mutation to evolve a population of solutions over successive generations. Individuals that are better suited to the environment survive and produce offspring with traits that are also well-adapted to the problem at hand. Both PSO and GA have been widely applied to various optimization problems in engineering, economics, and other fields, and each has its strengths and weaknesses in different problem domains. While PSO emphasizes the concept of collaboration and sharing of information among particles, GA relies on the principles of evolution and natural selection to converge toward optimal solutions.

Considering that both PSO and GA algorithms belong to the mathematical derivation is very clear and widely used in various types of research, this study will not list the mathematical derivation process of PSO and GA in detail (the process can be found in references^[29–35]), listed the selection

process of key hyperparameters as follows.

The key parameters of GA include population size, maximum number of iterations, selection probability, crossover probability, and mutation probability. Population size represents the number of individuals in each iteration. The maximum number of iterations signifies the maximum iteration steps. The selection, crossover, and mutation probabilities denote the likelihood of occurrence for selection, crossover, and mutation operations, respectively. The population size is fixed at 20, and the maximum number of iterations is set to 50. Drawing from relevant research, the selection, crossover, and mutation probabilities are chosen as 0.8, 0.2, and 0.2, respectively [29–32]. In contrast, the key parameters of PSO include population size, maximum number of iterations, inertia weight, and learning factors. The population size and maximum number of iterations are also set at 20 and 50, respectively. The inertia weight (w) determines the flight speed of particles, where a higher inertia weight allows particles to take longer strides for global exploration, while a smaller weight encourages finer local exploration. Referring to relevant studies, the inertia weight is set at 0.8 [33]. The learning factor C_1 represents the weight of a particle's own best position in influencing its next move, essentially pushing the particle towards its optimal position, while the learning factor C_2 represents the weight of the best position of other particles in steering a particle towards the collective optimal position. Based on pertinent research, the learning factors are set at 1.6 for C_1 and 1.8 for C_2 .

2.6 Squeeze-and-excitation (SE) block attention mechanism

The Squeeze-and-Excitation (SE) block is a structural component that prioritizes the channel dimension and incorporates a channel attention mechanism. This mechanism dynamically adjusts the relevance of each channel by assigning importance weights, enabling the enhancement or suppression of specific channels. Consequently, this attention mechanism facilitates the extraction of valuable features while disregarding irrelevant ones. The SE model is shown in Figure 3.

First, in Step 1 (Squeeze), features are compressed in spatial dimensions while the number of feature channels is maintained. Specifically, the module produces global pooling, which converts each two-dimensional feature channel to a real number (U_k) using the equation (9) shown below. M and N represent the spatial dimensions.

$$U_k = \frac{\sum_{m=1}^M \sum_{n=1}^N u_{m,n}}{M \times N}, m, n = 1, 2, \dots \quad (9)$$

Second, in Step 2 (Excitation), this module contains two fully connected layers, a Sigmoid function layer, and a ReLU function layer. S is the output of the Sigmoid function. *Sigmoid* and *ReLU* represent the two activation functions respectively. W_1 and W_2 are the corresponding parameters for these two fully connected layers. This step results in a lifting and then a reduction in the dimensionality of the features, as shown in equation (10).

$$S = \text{Sigmoid} [W_2 \times \text{ReLU}(W_1 \times U)] \quad (10)$$

Finally, in Step 3 (Reweight), the input features are weighted channel-by-channel to complete the weight reassignment of the original features.

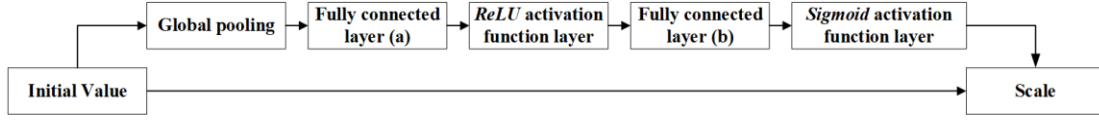


Figure 3 The step flow chart of the Squeeze-and-Excitation (SE) block

In the SE module, the *sigmoid* function will be used as the activation function, and the neurons of SE are set to 128 according to some recent studies ^[76–80] (It will be set to 128 or 256 generally). According to these studies, researchers proposed that a higher dimensionality would allow for more fine-grained attention but also increase the complexity and computational requirements of the model. In balancing these factors, 128 was chosen as a suitable compromise, offering sufficient granularity while keeping the computational demands within manageable limits ^[80]. Moreover, the SE module is usually added as a fully connected layer after the convolutional layer, which means SE is a single-layer structure without hidden layers.

2.7 Evaluation Indicators

The performance of the proposed method was evaluated using three indices: the coefficient of determination (R^2), the mean absolute percentage error (MAPE), and the coefficient-of-variation of root mean square error (CV-RMSE), computed using equations (11–13).

$$R^2(E_a, E_b) = 1 - \frac{\sum_{c=1}^n (E_{a,c} - E_{b,c})^2}{\sum_{c=1}^n (E_{b,c} - \bar{E})^2} \quad (11)$$

$$MAPE(E_a, E_b) = \frac{\sum_{c=1}^n \left| \frac{E_{a,c} - E_{b,c}}{E_{b,c}} \times 100\% \right|}{n} \quad (12)$$

$$CV - RMSE(E_a, E_b) = \frac{\sqrt{\frac{\sum_{c=1}^n (E_{a,c} - E_{b,c})^2}{n}}}{\bar{E}_b} \quad (13)$$

Where, E_a , and E_b is the prediction and actual energy consumption, kWh. In addition, to measure the improvement, PI is defined.

$$PI_{MAPE \text{ or } CV-RMSE} = \frac{M_0 - M_1}{M_0} \times 100\% \quad (14)$$

$$PI_R = \frac{M_1 - M_0}{M_0} \times 100\% \quad (15)$$

Where M_0 and M_1 are the values of the evaluation indices before and after development. For example, if comparing LSTM and CNN-LSTM, the M_0 is the performance of LSTM while the M_1 is the performance of CNN-LSTM. In the two equations, the positive values indicate performance improvement while negative values mean performance degradation.

In addition, it was important to evaluate the time required for implementation. It is the time of the whole algorithm realization, from the beginning of data import to the end of the output of evaluation index results, including the whole training and validation process, and its unit is second (s). In this study, the time consumption was recorded using the timing program provided by

MATLAB. The program was run on a specific computer (Lenovo, Tianyi 510 pro, CPU: i7-13700), with a consistent operator and other influencing factors such as parameter selection kept constant throughout the analysis.

3 Results and discussion

Possible parameters as input mainly include historical energy consumption parameters, outdoor meteorological parameters, indoor environment parameters, occupant-related parameters which are indirectly reflected by indoor CO₂ concentration, and time-related parameters. Equipment and envelope performance parameters and building form parameters which remain constant from hour to hour are not considered as inputs in this study. This is because, in the multivariate input and univariate output data-driven prediction, these parameters are basically kept constant at every hour and cannot be used to reflect the hourly changes in building energy consumption. They are mostly used for classification and feature extraction rather than for input.

Therefore, in this study, there are two kinds of inputs and six kinds of algorithms as follows.

Input A: Using all possible parameters as inputs, including time, historical energy consumption, indoor air temperature, indoor CO₂ concentration, indoor air humidity, indoor PM2.5 concentration, outdoor air relative humidity, outdoor air temperature, outdoor solar radiation intensity, longitudinal wind speed, latitudinal wind speed, cloudiness, and precipitation, referring to recent studies [38–52].

Input B: Selecting suitable parameters as inputs. The rule is if the correlation coefficient between this parameter and energy consumption is greater than 0.2, this parameter will be recognized as suitable input.

Algorithms A–F: LSTM, CNN-LSTM, LSTM-SE, CNN-LSTM-SE, LSTM-PSO, LSTM-GA. These algorithms included the most conventional LSTM, the proposed CNN-LSTM-SE, CNN-LSTM, and LSTM-SE used to validate that the combination of the CNN and SE modules is valuable, as well as LSTM-PSO and LSTM-GA which are commonly used recently.

Output: Building energy consumption.

3.1 Input selection and analyses

Forecasting the whole building energy consumption, AC energy consumption, lighting energy consumption, and plug-load energy consumption may have different suitable inputs. Therefore, the input selection will be categorized according to these four forecast scenes.

(1) For the whole building energy consumption forecast in Buildings A–G, the input selection results are shown in Figure 4. The conclusions are shown as follows:

(1-a) Historical energy consumption, time, and indoor CO₂ concentration are generally valid, which may be explained by the fact that these three factors can be used to reflect the intensity of occupant energy-use behaviors during a certain period of time, which can then be used as inputs for building energy consumption prediction.

(1-b) Indoor air temperature and humidity and outdoor air temperature are valid in most air-conditioned buildings (Buildings A-E), but outdoor air humidity is not sensitive to building energy consumption.

(1-c) Outdoor solar radiation intensity is valid in buildings where the lighting energy consumption exceeds a certain threshold (not valid in Buildings C and E) because this parameter is strongly correlated with lighting energy consumption.

(1-d) Cloudiness, wind speed, indoor PM2.5 concentration, and precipitation may affect the use of some equipment and are valid in a few scenarios. However, the correlation between these parameters and energy consumption is low. It can be proposed that cloudiness, wind speed, and indoor PM2.5 concentration may be largely ineffective.

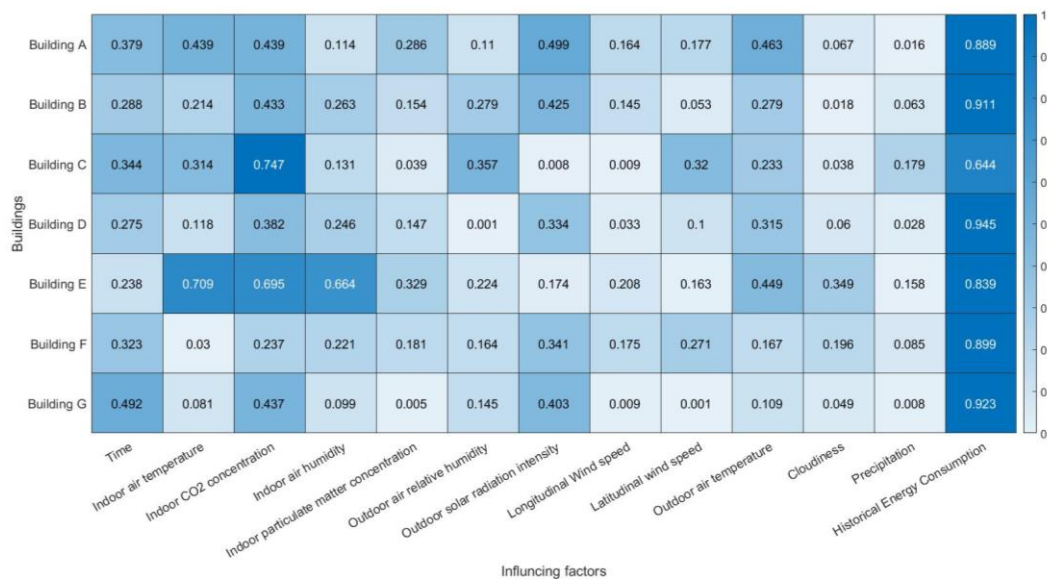


Figure 4 Correlation analysis results between the parameters and whole building energy consumption

(2) For the AC energy consumption forecast in Buildings A–E, the input selection results are shown in Figure 5. The conclusions are shown as follows:

(2-a) Historical energy consumption and indoor CO₂ concentration are generally valid, and the reason is similar to (1-a).

(2-b) Time, indoor air temperature and humidity, and outdoor air temperature and humidity are valid in most scenarios. The reason is that these parameters significantly affect air conditioning energy consumption

(2-c) Outdoor solar radiation intensity is applicable in some scenarios (Buildings A, C, and D), presumably because the outdoor solar radiation can significantly respond to hourly daily variations and yearly seasonal variations, making it similar to the conclusion in (2-b).

(2-d) Wind speed and indoor PM2.5 concentration are applicable in a small number of scenarios (Buildings A and E), and the common feature of Buildings A and E in terms of air conditioning energy consumption may be that Buildings A and E have small fluctuations in air

conditioning energy consumption. It may mean that occupants in Buildings A and E may not readily use or adjust the ACs, and it is only used to assist in the indoor environment creation in some scenarios such as extreme weather.

(2-e) Cloudiness and precipitation are only effective in some very specific scenarios with low correlation and do not have universal conclusions.

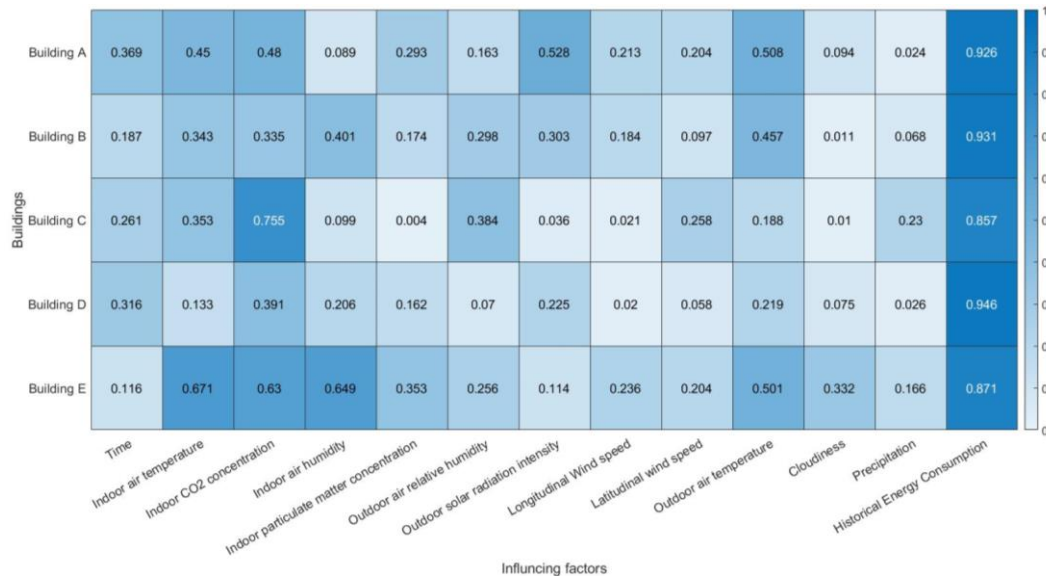


Figure 5 Correlation analysis results between the parameters and AC energy consumption

(3) For the lighting energy consumption forecast in Buildings A–G, the input selection results are shown in Figure 6. The conclusions are shown as follows:

(3-a) Historical energy consumption and indoor CO₂ concentration are generally valid, and the reason is similar to (1-a).

(3-b) Outdoor solar radiation intensity is generally valid. This may be due to the strong correlation between outdoor solar radiation and indoor illuminance, which significantly affects the occupant behaviors about using lighting equipment and thus significantly affects lighting energy consumption.

(3-c) Outdoor air temperature is universally valid, presumably because outdoor air temperature can respond significantly to seasonal variations, and thus indirectly to parameters such as hours of daylight, which are very strongly correlated with lighting energy consumption.

(3-d) Time, indoor air temperature and humidity, and outdoor air humidity are valid on some occasions, probably similar to the conclusion (3-b), and these parameters can respond more indirectly to seasonal changes.

(3-e) Cloudiness, wind speed, and indoor PM_{2.5} concentration were largely ineffective.

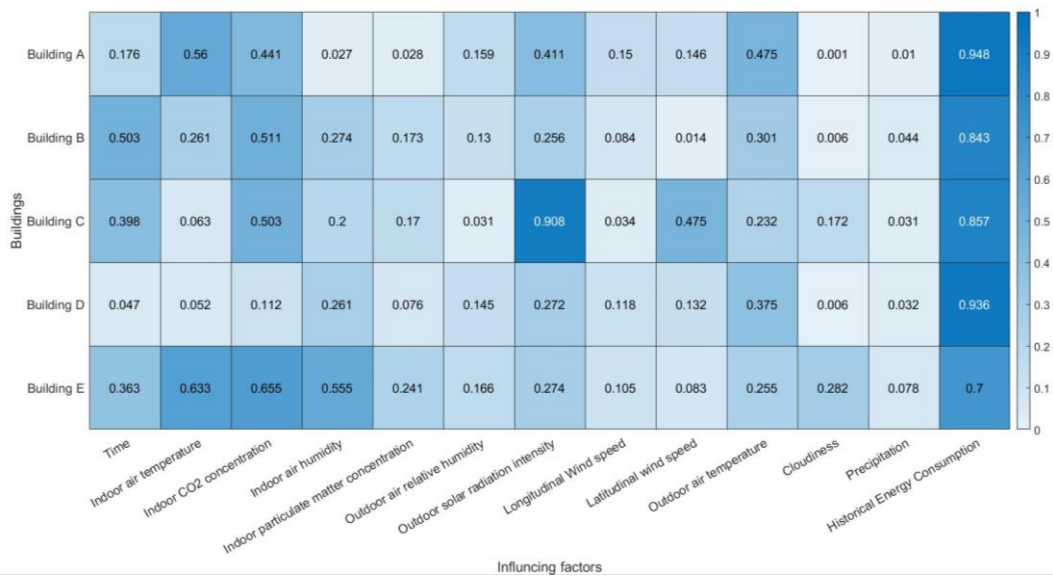


Figure 6 Correlation analysis results between the parameters and lighting energy consumption

(4) For the plug-load energy consumption forecast in Buildings A–G, the input selection results are shown in Figure 7. The conclusions are shown as follows:

(4-a) Historical energy consumption, time, and indoor CO₂ concentration are universally valid, and the reason is similar to (1-a).

(4-b) Indoor air temperature and humidity and outdoor air temperature and humidity are valid in some buildings, possibly because these parameters may affect the use of some cooling and heating equipment such as electric heaters.

(4-c) Similar to (1-c), outdoor solar radiation intensity is valid in buildings where the lighting energy share exceeds a certain threshold (it is not valid in Buildings C and E). Whereas almost all buildings have outlet branches that mix lights due to wiring or other reasons (similar scenes can be found in research [74–75]), it may cause this conclusion similar to (1-c).

(4-e) Cloudiness, wind speed, and indoor PM2.5 concentration are largely ineffective.

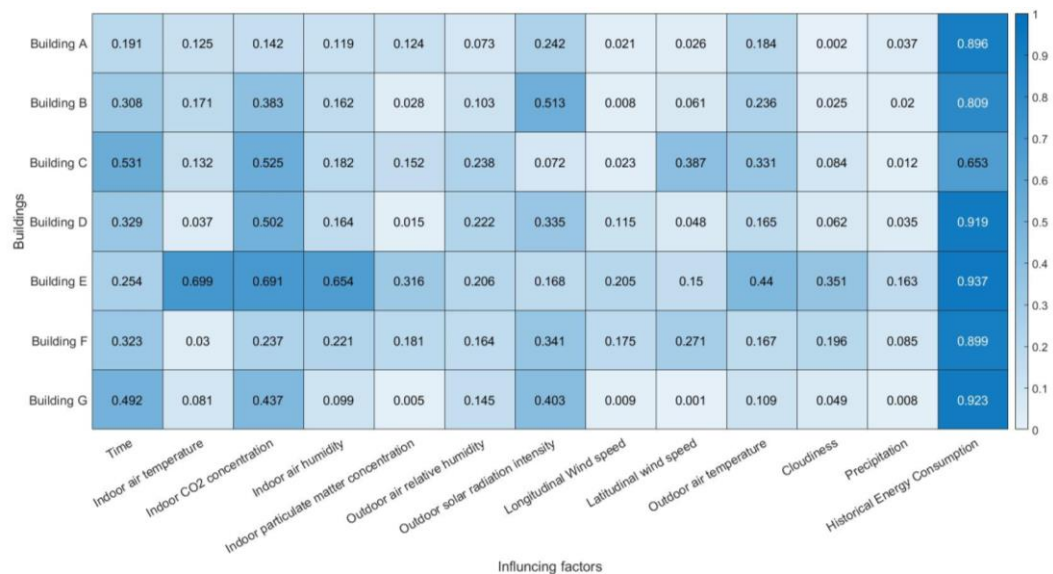


Figure 7 Correlation analysis results between the parameters and plug-load energy consumption

Considering the above conclusions, the suitable input parameters and corresponding scenarios in the different predictions can be listed in Table 3 below. These parameters can be divided into five categories:

(1) Generally valid (symbol: $\checkmark$): These parameters can be used as inputs in all scenarios without controversy;

(2) Valid in most scenarios (symbol: $\square$): These parameters can be used as inputs in most scenarios (e.g., solar radiation intensity for lighting energy consumption prediction, and outdoor temperature for air conditioning energy consumption prediction);

(3) Valid for some special scenarios (symbol: $\circ$): These parameters can be used as inputs in some special scenarios, which are shown in Table 3 below;

(4) Valid for a few non-special scenarios (symbol: $\times$): These parameters can be used as inputs in only a small number of scenarios, but these scenarios do not seem to have a clear common feature. Therefore, no clear conclusion can be given and it is recommended to validate them before each prediction;

(5) Generally invalid (symbol: $\times$): These parameters cannot be as inputs in most scenarios.

Table 3 The results of selecting suitable parameters for different predictions

	Building energy consumption	AC energy consumption	Lighting energy consumption	Plug-load energy consumption
Time	$\checkmark$	$\square$	$\times$	$\checkmark$
Historical energy consumption	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Indoor air temperature	$\circ(A)$	$\square$	$\times$	$\times$
Indoor CO ₂ concentration	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Indoor air humidity	$\circ(A)$	$\square$	$\times$	$\times$
Indoor PM2.5 concentration	$\times$	$\circ(C)$	$\times$	$\times$
Outdoor air relative humidity	$\times$	$\square$	$\times$	$\times$
Outdoor air temperature	$\circ(A)$	$\square$	$\square$	$\times$
Outdoor solar radiation intensity	$\circ(B)$	$\square$	$\checkmark$	$\circ(D)$
Longitudinal Wind speed	$\times$	$\circ(C)$	$\times$	$\times$
Latitudinal wind speed	$\times$	$\circ(C)$	$\times$	$\times$
Cloudiness	$\times$	$\times$	$\times$	$\times$
Precipitation	$\times$	$\times$	$\times$	$\times$

Note: Scenario A: There are air conditioners in the building (Buildings A-E in this study). Scenario B: The proportion of lighting energy consumption to total energy consumption in the building exceeds a certain threshold (Buildings A, B, D, F, and G in this study). Scenario C: Building air conditioning energy consumption fluctuates less (Buildings A, and E in this study). The reason for this is hypothesized to be that occupants in this type of building may not readily use or adjust the air conditioning, and it is only used to assist in creating an indoor environment in some scenarios such as extreme weather. Scenario D: The proportion of lighting energy consumption to total energy consumption in the building exceeds a certain threshold (Buildings A, B, D, F, and G in this study). This scenario is similar to Scenario B, presumably because this parameter is more strongly correlated with lighting energy

consumption. The similarity occurs because almost all buildings have outlet branches that mix lights due to wiring or other reasons.

3.2 Building energy consumption prediction with different inputs and algorithms

Taking energy consumption prediction for the whole building as the sample, eighty-four sets of data simulation experiments consisting of seven different buildings, two kinds of inputs, and six kinds of algorithms are cross-composed for validation. And selecting these four typical cases for comparative analyses:

Case a: Using LSTM as the algorithm and all parameters as inputs;

Case b: Using LSTM-PSO as the algorithm and all parameters as inputs;

Case c: Using LSTM-GA as the algorithm and all parameters as inputs;

Case d: Using CNN-LSTM-SE as the algorithm and selecting suitable parameters as inputs.

It can be calculated that Case d decreased 1.76%–12.82% MAPE (Average: 5.61%), increased 3.45%–13.28% R^2 (Average: 7.26%), decreased 1.98%–8.42% CV-RMSE (Average: 5.69%), and decreased 822.53–5055.57s (Average: 2771.65s) compared to Cases b-c on average. The results of these four selected cases are shown in Figure 8.

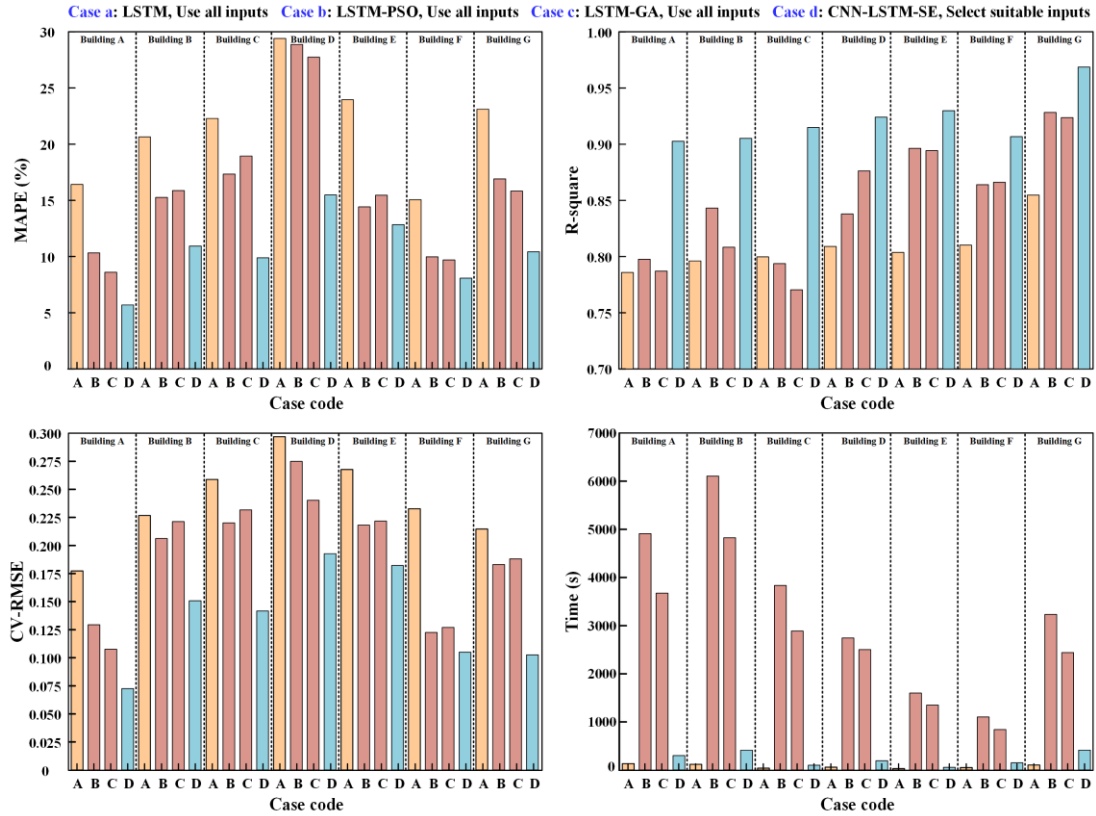


Figure 8 The results of building energy consumption prediction with different inputs and algorithms

All results of eighty-four simulation experiments are shown in Appendix A. For two kinds of inputs and six kinds of algorithms (There are twelve combinations after the permutation), the combination of using suitable parameters as inputs and CNN-LSTM-SE as algorithms may have the best performance which balances the time consumption and prediction precision better. The

performance improvement ranges are shown in Figure 9, compared to the other eleven combinations on average in different buildings. The eleven performance change data (All changes are performance improvement and no performance degradation has occurred) are plotted as a box-line diagram. Due to seven buildings and three evaluation indicators, Figure 9 includes twenty-one box-line structures.

For actual values, the combinations employing suitable parameters as inputs and CNN-LSTM-SE as algorithms can improve the prediction performance by decreasing 1.15%–65.32% of MAPE (Average: 28.87%), increasing 0.27%–18.74% of R^2 (Average: 7.61%), and decreasing 0.64 %–59.12% of CV-RMSE (Average: 26.63%) in different buildings. It will reduce by 2712.20s compared to the combinations with LSTM-PSO or LSTM-GA and increase by only 62.82s compared to combinations with LSTM, CNN-LSTM, or LSTM-SE on average.

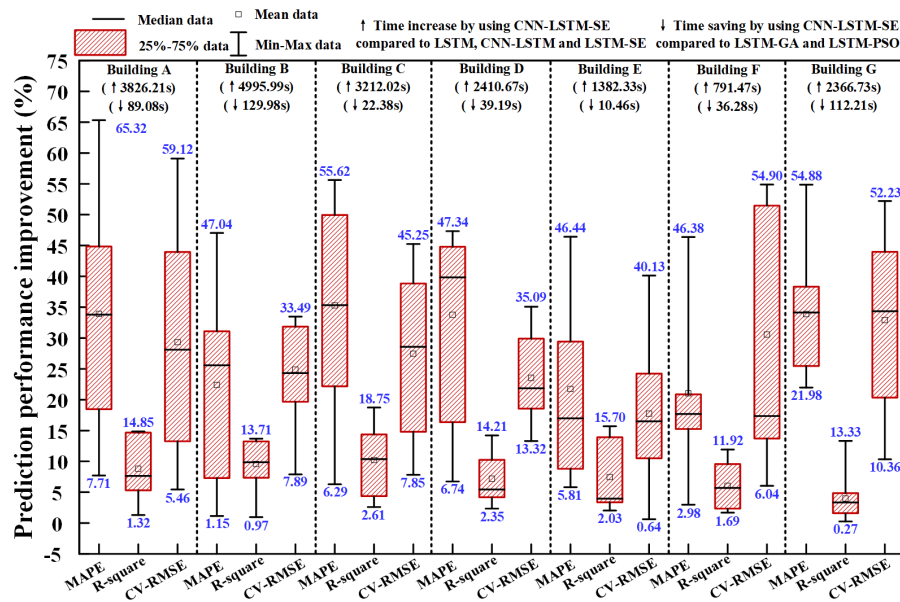


Figure 9 Performance improvement by selecting suitable parameters as inputs and using CNN-LSTM-SE as the training algorithm simultaneously

The performance improvement in Figure 9 is from two sources. The performance improvement from enhancing inputs alone and enhancing the training algorithm alone will be discussed as follows.

When discussing the comparison of prediction performance change resulting from selecting the suitable parameters as inputs or not (using all parameters as inputs), the training algorithms will remain fixed. The improvement values are 3.56%–25.56% (Average: 13.94%) decrease of MAPE, 2.34%–5.20% (Average: 3.36%) increase of R^2 , and 4.26%–18.25% (Average: 11.20%) decrease of CV-RMSE. Moreover, it can decrease time consumption by 25.34s–128.46s (Average: 50.56s). The performance improvement ranges from inputs are shown in Figure 10, with graphical composition illustrated similarly to Figure 9.

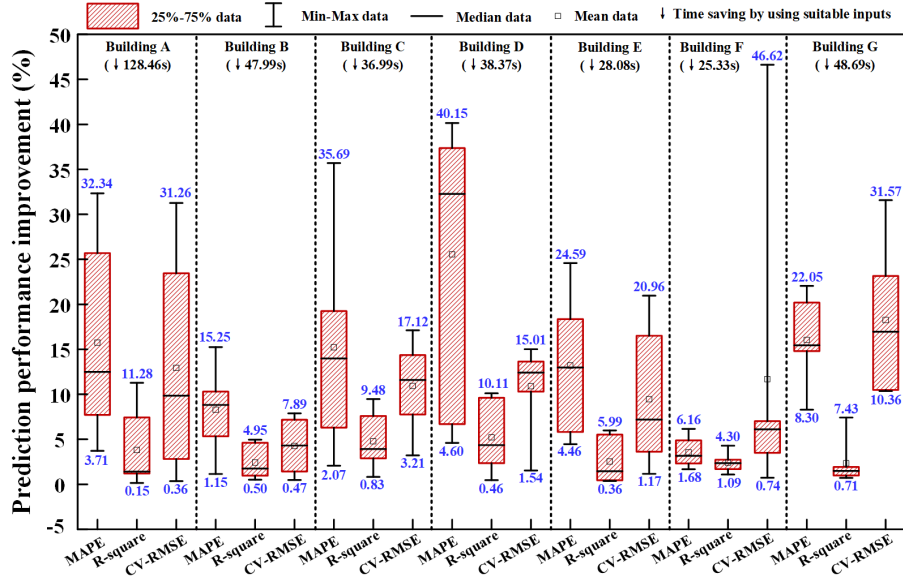


Figure 10 Performance improvement results by using suitable parameters as inputs

When discussing the comparison of prediction performance change resulting from using CNN-LSTM-SE as the training algorithm or not (using other algorithms), the inputs will remain fixed. The improvement values are 15.79 %–32.01% (Average: 21.44%) decrease of MAPE, 2.69%–9.31% (Average: 6.25%) increase of R^2 , and 15.68%–29.23% (Average: 21.82 %) decrease of CV-RMSE. Moreover, it can decrease time consumption by 781.40s–4988.94s (Average: 2696.89s) compared to LSTM-PSO or LSTM-GA, and increase it by 20.69s–161.04s (Average: 91.30s) compared to CNN-LSTM, or LSTM-SE on average. It is also important to note that since the prediction performance of LSTM without any optimization module is so poor, the comparison did not include LSTM and only included LSTM-SE, LSTM-PSO, LSTM-GA, and CNN-LSTM. The performance improvement ranges from algorithms are shown in Figure 11, with graphical composition illustrated similarly to Figure 9.

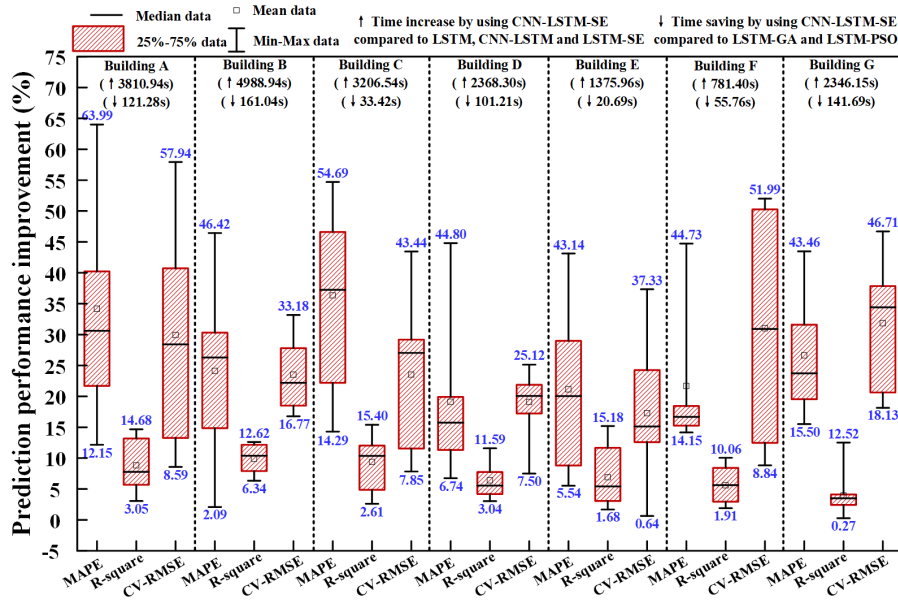


Figure 11 Performance improvement results by using CNN-LSTM-SE as the training algorithm alone

The detailed data about prediction performance improvement by selecting suitable inputs and using CNN-LSTM-SE in different buildings will be shown in Table 4.

Table 4 Prediction performance improvement from selecting suitable inputs and using CNN-LSTM-SE on average

Building	Prediction performance improvement by selecting suitable parameters as inputs				Prediction performance improvement by using CNN-LSTM-SE as the training algorithm				
	MAPE	R ²	CV-RMSE	Time	MAPE	R ²	CV-RMSE	Time	Time
	decrease	increase	decrease	decrease (s)	decrease	increase	decrease	Increase (s)	decrease (s)
A	15.738%	3.815%	12.927%	128.4628	26.912%	7.531%	23.072%	124.2820	3810.9428
B	8.278%	2.427%	4.264%	47.9917	19.231%	9.308%	21.752%	161.0357	4988.9405
C	15.209%	4.771%	10.938%	36.9945	32.006%	8.645%	19.413%	33.4222	3206.5365
D	25.556%	5.204%	10.887%	38.3702	15.790%	6.006%	17.763%	102.2108	2368.2990
E	13.197%	2.540%	9.441%	28.0787	17.403%	5.058%	15.677%	20.6880	1375.9600
F	3.562%	2.416%	11.680%	25.3373	16.154%	4.531%	25.833%	55.7602	781.4018
G	16.041%	2.337%	18.249%	48.6860	22.583%	2.686%	29.230%	141.6883	2346.1520

Note: 1. MAPE decrease, R² increase, and CV-RMSE decrease represent the percentage of performance improvement (in %), while Time decrease and Time increase represent the value of time increase or decrease (in s). 2. Time increase represents the increase in time consumption by using CNN-LSTM-SE, mainly comparing the three algorithms: LSTM, CNN-LSTM, and LSTM-SE, while time decrease represents the decrease in time consumption using CNN-LSTM-SE, mainly comparing the two algorithms: LSTM-PSO, and LSTM-GA.

4 Limits and future directions

(1) In this study, the correlation coefficient method was used to select the suitable inputs with the judgment threshold of correlation being 0.2. Although many studies have endorsed the selection of this threshold, it is assumed that the correlation coefficient between parameter A and energy consumption is 0.199, while it is 0.201 for parameter B, and there does not appear to be a significant difference between the effects of parameters A and B on energy consumption. However, they will be completely differentiated in this study. Although the parameters in this study do not seem to

present this special case, this phenomenon needs to be carefully considered in subsequent studies. That is, it may be more reasonable and scientific if the judgment of the correlation threshold is segmented and fuzzy rather than a fixed value.

(2) More different kinds of buildings may be needed. In particular, in this study, some scenarios require specific data to be constrained (e.g., AC energy consumption share of total energy consumption, CV and IQR of energy consumption distribution, etc.), which means that more buildings with different types and located in different climate zones are needed for validation.

(3) In the context of algorithms, this study is only preliminary in its consideration of the attention mechanisms, because attention mechanisms encompass a variety of branches, such as Channel attention, Spatial attention, and Branch attention (SE attention, employed in this study, constitutes a kind of channel attention). Additionally, some other recent optimizations were not considered in this study, especially some recent swarm intelligence algorithms such as the sparrow search algorithm (SSA). These swarm intelligent optimization algorithms may have been improved the computationally cumbersome problem compared to the PSO. Therefore, more studies for more attention mechanisms and swarm intelligent optimization algorithms need to be considered in the future.

5 Conclusions

This study presents a comprehensive method for enhancing building energy consumption prediction, which includes selecting appropriate inputs and improving algorithms for different forecast targets including forecasting energy consumption for whole building energy consumption, AC systems, and plug loads.

Regarding selecting appropriate input, it is crucial to use suitable inputs for accurate energy consumption prediction. In this study, many parameters were discussed for their suitability as inputs. Slightly different from other studies, this study confirms that indoor carbon dioxide concentration performs better in various scenarios. Moreover, the remaining parameters such as time, historical energy consumption, and meteorological parameters in which scenarios are applicable are also provided. Compared to using all possible parameters as inputs, selecting suitable parameters as inputs can improve the prediction performance with 3.56%–25.56% (Average: 13.94%) decrease of MAPE, 2.34%–5.20% (Average: 3.36%) increase of R^2 , and 4.26%–18.25% (Average: 11.20%) decrease of CV-RMSE. Moreover, it can decrease time consumption by 25.34s–128.46s (Average: 50.56s). These results suggest the necessity and importance of selecting suitable inputs when forecasting building energy consumption.

Regarding improving algorithms, this study integrates CNN for optimizing the data processing process and introduces the SE attention mechanism (SE) for optimizing the hyperparameter optimization process. Therefore, compared with CNN optimization alone, CNN-SE can have one

more enhancement path which is optimizing hyperparameters such as threshold and weights of the neural network. From the actual data experiments, the CNN-SE optimization will improve the building energy consumption prediction in most scenes more than CNN optimization. Compared to the other four algorithms including CNN-LSTM, LSTM-SE, LSTM-PSO, and LSTM-GA, using CNN-LSTM-SE as the training algorithm can improve the prediction performance with 15.79 %–32.01% (Average: 21.44%) decrease of MAPE, 2.69%–9.31% (Average: 6.25%) increase of R^2 , and 15.68%–29.23% (Average: 21.82 %) decrease of CV-RMSE. Moreover, it can decrease time consumption by 781.40s–4988.94s (Average: 2696.89s) compared to LSTM-PSO or LSTM-GA, and increase it by 20.69s–161.04s (Average: 91.30s) compared to CNN-LSTM, or LSTM-SE on average.

In summary, in the twelve combinations (two kinds of inputs: selecting suitable inputs or not, six kinds of algorithms: LSTM, CNN-LSTM, LSTM-SE, LSTM-PSO, LSTM-GA, CNN-LSTM-SE), the combination of selecting suitable parameters as inputs and using CNN-LSTM-SE as algorithms may have the best performance which balances the time consumption and prediction precision better. This combination can improve the prediction performance by decreasing 1.15%–65.32% of MAPE (Average: 28.87%), increasing 0.27%–18.74% of R^2 (Average: 7.61%), and decreasing 0.64 %–59.12% of CV-RMSE (Average: 26.63%) in different buildings compared to the other eleven combinations. It will reduce by 2712.20s compared to the combinations with LSTM-PSO or LSTM-GA and increase by only 62.82s compared to combinations with LSTM, CNN-LSTM, or LSTM-SE on average.

It is important to note that importantly this study demonstrates the significance of preferred inputs. However, the parameters selected in this study are not necessarily fully applicable to other studies. It is difficult to draw universal conclusions about the applicability or non-applicability of certain parameters to different research subjects, and the selection should be based on the actual situation (building type, building energy consumption characteristics, building operation time, etc.).

Author Statement

Chengyu Zhang: Conceptualization, Methodology, Validation, Investigation, Formal analysis, Visualization, Writing - Original Draft, Writing - review & editing

Zhiwen Luo: Writing - review & editing

Yacine Rezgui: Writing - review & editing

Tianyi Zhao: Conceptualization, Resources, Data Curation, Supervision, Project administration, Funding acquisition

Acknowledgment

The work described in this paper was supported by the National Natural Science Foundation of

China (Grant No. 52078096), the Natural Science Foundation Joint Fund of Liaoning Province (Grant No. 2023-MSBA-023), 2023 Dalian University of Technology-Cardiff University Cooperation and Exchange Foundation Project, 2023 International Exchange Foundation Project of "Co-Creation of Excellence Program" from Dalian University of Technology (Grant: DUTIO-ZG-202307), the Key Project of DUT for International Students Studying and Researching in China: Innovation and Practice of Talent Cultivation Model in the Field of Smart Buildings for the "Belt and Road" Initiative (Grant No. 1103-82120001).

Declaration of interests

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.

Appendix A: Results of building energy consumption prediction

The results of building energy consumption prediction employing different inputs and training algorithms are shown in Table A-1 as follows.

Table A-1 Results of building energy consumption prediction

Building	Algorithms	Input selection	MAPE	R ²	CV-RMSE	Time (s)
A	LSTM	Use all inputs	16.42%	0.7859	0.1774	132.4840
A	CNN-LSTM	Use all inputs	8.72%	0.8259	0.0839	265.9490
A	LSTM-SE	Use all inputs	9.24%	0.8430	0.1176	217.9380
A	CNN-LSTM-SE	Use all inputs	6.17%	0.8908	0.0767	331.9710
A	LSTM-PSO	Use all inputs	10.32%	0.7976	0.1294	4911.1630
A	LSTM-GA	Use all inputs	8.60%	0.7871	0.1076	3675.1410
A	LSTM	Select suitable inputs	15.81%	0.7871	0.1724	77.6510
A	CNN-LSTM	Select suitable inputs	6.48%	0.8384	0.0836	250.7250
A	LSTM-SE	Select suitable inputs	8.36%	0.8530	0.1009	209.7790
A	CNN-LSTM-SE	Select suitable inputs	5.69%	0.9026	0.0725	301.4350
A	LSTM-PSO	Select suitable inputs	6.98%	0.8569	0.0889	4740.1990
A	LSTM-GA	Select suitable inputs	7.27%	0.8759	0.0823	3184.0800
B	LSTM	Use all inputs	20.65%	0.7961	0.2267	119.1730
B	CNN-LSTM	Use all inputs	11.80%	0.7993	0.1967	350.6420
B	LSTM-SE	Use all inputs	15.16%	0.7993	0.1993	320.2190
B	CNN-LSTM-SE	Use all inputs	11.06%	0.8966	0.1637	425.5050
B	LSTM-PSO	Use all inputs	15.27%	0.8431	0.2062	6107.9720
B	LSTM-GA	Use all inputs	15.88%	0.8083	0.2212	4825.9920
B	LSTM	Select suitable inputs	18.52%	0.8088	0.2257	108.9120
B	CNN-LSTM	Select suitable inputs	11.17%	0.8389	0.1877	326.5020
B	LSTM-SE	Select suitable inputs	12.85%	0.8363	0.1850	319.0920
B	CNN-LSTM-SE	Select suitable inputs	10.94%	0.9053	0.1508	411.4130
B	LSTM-PSO	Select suitable inputs	13.71%	0.8474	0.1978	5932.3410
B	LSTM-GA	Select suitable inputs	14.70%	0.8239	0.2181	4763.2930
C	LSTM	Use all inputs	22.29%	0.7998	0.2588	39.2470
C	CNN-LSTM	Use all inputs	13.34%	0.8479	0.1803	95.4170
C	LSTM-SE	Use all inputs	19.77%	0.8144	0.1855	96.4710
C	CNN-LSTM-SE	Use all inputs	10.56%	0.8891	0.1641	112.7710
C	LSTM-PSO	Use all inputs	17.35%	0.7937	0.2201	3835.5100
C	LSTM-GA	Use all inputs	18.94%	0.7705	0.2317	2888.0900
C	LSTM	Select suitable inputs	21.83%	0.8064	0.2505	39.1770
C	CNN-LSTM	Select suitable inputs	11.54%	0.8766	0.1663	94.5490
C	LSTM-SE	Select suitable inputs	12.71%	0.8916	0.1537	78.3220
C	CNN-LSTM-SE	Select suitable inputs	9.89%	0.9149	0.1417	101.8010
C	LSTM-PSO	Select suitable inputs	14.84%	0.8290	0.1991	3649.7400
C	LSTM-GA	Select suitable inputs	15.30%	0.8289	0.1984	2881.9500
D	LSTM	Use all inputs	29.40%	0.8091	0.2969	59.7430
D	CNN-LSTM	Use all inputs	25.74%	0.8104	0.2631	167.8730
D	LSTM-SE	Use all inputs	26.56%	0.8449	0.2824	204.4700

D	CNN-LSTM-SE	Use all inputs	23.55%	0.9028	0.2223	277.4440
D	LSTM-PSO	Use all inputs	28.87%	0.8380	0.2749	2744.1620
D	LSTM-GA	Use all inputs	27.74%	0.8762	0.2403	2501.8460
D	LSTM	Select suitable inputs	28.05%	0.8868	0.2523	56.8700
D	CNN-LSTM	Select suitable inputs	24.02%	0.8923	0.2328	165.1950
D	LSTM-SE	Select suitable inputs	18.52%	0.8855	0.2439	143.0460
D	CNN-LSTM-SE	Select suitable inputs	15.48%	0.9240	0.1927	192.7100
D	LSTM-PSO	Select suitable inputs	18.08%	0.8707	0.2466	2665.8760
D	LSTM-GA	Select suitable inputs	16.60%	0.8802	0.2366	2501.6200
E	LSTM	Use all inputs	23.96%	0.8037	0.2677	29.4070
E	CNN-LSTM	Use all inputs	18.19%	0.8162	0.2035	53.9920
E	LSTM-SE	Use all inputs	21.91%	0.8484	0.3042	48.9170
E	CNN-LSTM-SE	Use all inputs	13.63%	0.9113	0.1906	69.1990
E	LSTM-PSO	Use all inputs	14.43%	0.8963	0.2181	1599.6290
E	LSTM-GA	Use all inputs	15.46%	0.8943	0.2217	1349.9490
E	LSTM	Select suitable inputs	18.07%	0.8073	0.2235	28.7620
E	CNN-LSTM	Select suitable inputs	15.10%	0.8612	0.1833	53.7840
E	LSTM-SE	Select suitable inputs	17.89%	0.8992	0.2404	38.0110
E	CNN-LSTM-SE	Select suitable inputs	12.83%	0.9299	0.1821	56.4680
E	LSTM-PSO	Select suitable inputs	13.78%	0.8995	0.2156	1463.7340
E	LSTM-GA	Select suitable inputs	14.07%	0.9023	0.2136	1341.8620
F	LSTM	Use all inputs	15.06%	0.8102	0.2327	50.8460
F	CNN-LSTM	Use all inputs	9.86%	0.8450	0.2246	141.6340
F	LSTM-SE	Use all inputs	10.21%	0.8225	0.2128	123.3390
F	CNN-LSTM-SE	Use all inputs	8.32%	0.8917	0.1117	169.8310
F	LSTM-PSO	Use all inputs	9.98%	0.8640	0.1225	1102.4070
F	LSTM-GA	Use all inputs	9.70%	0.8661	0.1270	842.0260
F	LSTM	Select suitable inputs	14.13%	0.8276	0.2164	50.1950
F	CNN-LSTM	Select suitable inputs	9.53%	0.8542	0.1199	137.8550
F	LSTM-SE	Select suitable inputs	9.71%	0.8578	0.1997	120.1180
F	CNN-LSTM-SE	Select suitable inputs	8.08%	0.9068	0.1050	149.6850
F	LSTM-PSO	Select suitable inputs	9.81%	0.8859	0.1216	1013.0170
F	LSTM-GA	Select suitable inputs	9.47%	0.8898	0.1226	807.1890
G	LSTM	Use all inputs	23.11%	0.8547	0.2147	105.8480
G	CNN-LSTM	Use all inputs	16.42%	0.9278	0.1745	346.6420
G	LSTM-SE	Use all inputs	16.63%	0.9533	0.1441	437.5790
G	CNN-LSTM-SE	Use all inputs	13.38%	0.9618	0.1144	454.7150
G	LSTM-PSO	Use all inputs	16.91%	0.9282	0.1830	3232.2750
G	LSTM-GA	Use all inputs	15.83%	0.9236	0.1881	2441.0360
G	LSTM	Select suitable inputs	18.45%	0.9182	0.1650	103.6210
G	CNN-LSTM	Select suitable inputs	13.99%	0.9457	0.1562	335.8760
G	LSTM-SE	Select suitable inputs	15.25%	0.9660	0.1252	325.1020
G	CNN-LSTM-SE	Select suitable inputs	10.43%	0.9686	0.1025	413.5510
G	LSTM-PSO	Select suitable inputs	14.32%	0.9373	0.1449	3123.2260

G	LSTM-GA	Select suitable inputs	13.37%	0.9389	0.1287	2324.6030
---	---------	------------------------	--------	--------	--------	-----------

References

- [1] A Streltsov, J M Malof, B Huang, K Bradbury. Estimating residential building energy consumption using overhead imagery. *Applied Energy*, 2020, 280, 116018. <https://doi.org/10.1016/j.apenergy.2020.116018>
- [2] X Cao, X Dai, J Liu. Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade. *Energy and Buildings*, 2016, 128, 198-213. <https://doi.org/10.1016/j.enbuild.2016.06.089>
- [3] China Association of Building Energy Efficiency. 2022 China Urban and Rural Construction Carbon Emission Series Research Report. 2023.01.04. In Chinese. From: <https://www.cabee.org/site/content/24420.html>
- [4] Li L, Zhang S. Techno-economic and environmental assessment of multiple distributed energy systems coordination under centralized and decentralized framework. *Sustainable Cities and Society*, 2021, 72, 103076. <https://doi.org/10.1016/j.scs.2021.103076>
- [5] Qi S, Zhou C, Li K et al. The impact of a carbon trading pilot policy on the low-carbon international competitiveness of industry in China: An empirical analysis based on a DDD model. *Journal of Cleaner Production*, 2021, 281, 125361. <https://doi.org/10.1016/j.jclepro.2020.125361>
- [6] Jiang H, Jiang P, Wang Det al. Can smart city construction facilitate green total factor productivity? A quasi-natural experiment based on China's pilot smart city. *Sustainable Cities and Society*, 2021, 69, 102809. <https://doi.org/10.1016/j.scs.2021.102809>
- [7] Clabeaux R., Carbajales-Dale M., Ladner D., et al. Assessing the carbon footprint of a university campus using a life cycle assessment approach. *Journal of Cleaner Production*. 2020, 273, 122600. <https://doi.org/10.1016/j.jclepro.2020.122600>
- [8] Maltais, L., & Gosselin, L. Forecasting of short-term lighting and plug load electricity consumption in single residential units: Development and assessment of data-driven models for different horizons. *Applied Energy*. 2022, 307, 118229. <https://doi.org/10.1016/j.apenergy.2021.118229>
- [9] Fischer, D., Bernhardt, J., Madani, H., & Wittwer, C. Comparison of control approaches for variable speed air source heat pumps considering time variable electricity prices and PV. *Applied Energy*. 2017, 204, 93-105. <https://doi.org/10.1016/j.apenergy.2017.06.110>
- [10] A., S., M., F., M., M., A., A.F., J., R.S., & R., P.L. Forecast-Based Energy Management for Domestic PV-Battery Systems: A U.K. Case Study. *IEEE Access*. 2021, 9, 58953-58965. <https://doi.org/10.1109/ACCESS.2021.3072961>
- [11] D Yan, T Hong, B Dong, A Mahdavi, S D Oca, I Gaetani, X Feng. IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings. *Energy and Buildings*, 2017, 156, 258-270. <https://doi.org/10.1016/j.enbuild.2017.09.084>
- [12] H Sha, P Xu, C Yan, Y Ji, K Zhou, F Chen. Development of a key-variable-based parallel HVAC energy predictive model. *Building Simulation*, 2022, 15(7), 1193-1208. <https://doi.org/10.1007/s12273-021-0885-0>
- [13] Amasyali, K., & El-Gohary, N.M. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*. 2018, 81, 1192-1205. <https://doi.org/10.1016/j.rser.2017.04.095>

- [14] M H Shamsi, U Ali, E Mangina, J O Donnell. Feature assessment frameworks to evaluate reduced-order grey-box building energy models. *Applied Energy*, 2021, 298, 117174. <https://doi.org/10.1016/j.apenergy.2021.117174>
- [15] Y Li, Z O'Neill, L Zhang, J Chen, P Im, J Degraw. Grey-box modeling and application for building energy simulations - A critical review. *Renewable and Sustainable Energy Reviews*, 2021, 146, 111174. <https://doi.org/10.1016/j.rser.2021.111174>
- [16] China Meteorological Administration. China Meteorological Science Data Center. From: <http://data.cma.cn>
- [17] A Franco, F Leccese. Measurement of CO₂ concentration for occupancy estimation in educational buildings with energy efficiency purposes. *Journal of Building Engineering*, 2020, 32, 101714. <https://doi.org/10.1016/j.jobe.2020.101714>
- [18] M S Zuraimi, A Pantazaras, K A Chaturvedi, J J Yang, K W Tham, S E Lee. Predicting occupancy counts using physical and statistical Co₂-based modeling methodologies. *Building and Environment*, 2017, 123, 517-528. <https://doi.org/10.1016/j.buildenv.2017.07.027>
- [19] H Rahman, H Han. Bayesian estimation of occupancy distribution in a multi-room office building based on CO₂ concentrations. *Building Simulation*, 2018, 11(3), 575-583. <https://doi.org/10.1007/s12273-017-0413-9>
- [20] A Szczurek, M Maciejewska, T Pietrucha. Occupancy determination based on time series of CO₂ concentration, temperature and relative humidity. *Energy and Buildings*, 2017, 147, 142-154. <https://doi.org/10.1016/j.enbuild.2017.04.080>
- [21] K Sun, Q Zhao, J Zou. A review of building occupancy measurement systems. *Energy and Buildings*, 2020, 216, 109965. <https://doi.org/10.1016/j.enbuild.2020.109965>
- [22] G Xianliang, X Jingchao, L Zhiwen, L Jiaping. Analysis to energy consumption characteristics and influencing factors of terminal building based on airport operating data. *Sustainable Energy Technologies and Assessments*, 2021, 44, 101034. <https://doi.org/10.1016/j.seta.2021.101034>
- [23] L Yuan, Y Ruan, G Yang, F Feng, Z Li. Analysis of Factors Influencing the Energy Consumption of Government Office Buildings in Qingdao. *Energy Procedia*, 2016, 104, 263-268. <https://doi.org/10.1016/j.egypro.2016.12.045>
- [24] S Zhan, B Dong, A Chong. Improving energy flexibility and PV self-consumption for a tropical net zero energy office building. *Energy and Buildings*, 2023, 278, 112606. <https://doi.org/10.1016/j.enbuild.2022.112606>
- [25] M Khalil, A S Mcgough, Z Pourmirza, M Pazhoohesh, S Walker. Machine Learning, Deep Learning and Statistical Analysis for forecasting building energy consumption — A systematic review. *Engineering Applications of Artificial Intelligence*, 2022, 115, 105287. <https://doi.org/10.1016/j.engappai.2022.105287>
- [26] F Yao, W Zhou, M A Ghamdi, Y Song, W Zhao. An integrated D-CNN-LSTM approach for short-term heat demand prediction in district heating systems. *Energy Reports*, 2022, 8, 98-107. <https://doi.org/10.1016/j.egyr.2022.08.087>
- [27] R Rick, L Berton. Energy forecasting model based on CNN-LSTM-AE for many time series with unequal lengths. *Engineering Applications of Artificial Intelligence*, 2022, 113, 104998. <https://doi.org/10.1016/j.engappai.2022.104998>

- [28] M Yang, M K Lim, Y Qu, X Li, D Ni. Deep neural networks with L1 and L2 regularization for high dimensional corporate credit risk prediction. *Expert Systems with Applications*, 2023, 213, 118873. <https://doi.org/10.1016/j.eswa.2022.118873>
- [29] X Luo, L O Oyedele. A self-adaptive deep learning model for building electricity load prediction with moving horizon. *Machine Learning with Applications*, 2022, 7, 100257. <https://doi.org/10.1016/j.mlwa.2022.100257>
- [30] X J Luo, L O Oyedele. Forecasting building energy consumption: Adaptive long-short term memory neural networks driven by genetic algorithm. *Advanced Engineering Informatics*, 2021, 50, 101357. <https://doi.org/10.1016/j.aei.2021.101357>
- [31] X J Luo, L O Oyedele, A O Ajayi, O O Akinade, J M D Delgado, H A Owolabi, A Ahmed. Genetic algorithm-determined deep feedforward neural network architecture for predicting electricity consumption in real buildings. *Energy and AI*, 2020, 2, 100015. <https://doi.org/10.1016/j.egyai.2020.100015>
- [32] X J Luo, L O Oyedele, A O Ajayi, O O Akinade, H A Owolabi, A Ahmed. Feature extraction and genetic algorithm enhanced adaptive deep neural network for energy consumption prediction in buildings. *Renewable and Sustainable Energy Reviews*, 2020, 131, 109980. <https://doi.org/10.1016/j.rser.2020.109980>
- [33] Y Cao, G Liu, J Sun, D P Bavarisetti, G Xiao. PSO-Stacking improved ensemble model for campus building energy consumption forecasting based on priority feature selection. *Journal of Building Engineering*, 2023, 72, 106589. <https://doi.org/10.1016/j.jobe.2023.106589>
- [34] A G Gad. Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review. *Archives of Computational Methods in Engineering*, 2022, 29(5), 2531-2561. <https://doi.org/10.1007/s11831-021-09694-4>
- [35] G Zhang, Y Ge, X Pan, M Sadat Afsharzadeh, M Ghalandari. Optimization of energy consumption of a green building using PSO-SVM algorithm. *Sustainable Energy Technologies and Assessments*, 2022, 53, 102667. <https://doi.org/10.1016/j.seta.2022.102667>
- [36] B Kuang, C Schelly, G Ou, M Sahraei-Ardakani, S Tiwari, J Chen. Data-driven analysis of influential factors on residential energy end-use in the US. *Journal of Building Engineering*, 2023, 75, 106947. <https://doi.org/10.1016/j.jobe.2023.106947>
- [37] Y Mo, D Zhao. Effective factors for residential building energy modeling using feature engineering. *Journal of Building Engineering*, 2021, 44, 102891. <https://doi.org/10.1016/j.jobe.2021.102891>
- [38] Y Peng, Z Nagy, A Schlüter. Temperature-preference learning with neural networks for occupant-centric building indoor climate controls. *Building and Environment*, 2019, 154, 296-308. <https://doi.org/10.1016/j.buildenv.2019.01.036>
- [39] S Mun, Y Kwak, J Huh. A case-centered behavior analysis and operation prediction of AC use in residential buildings. *Energy and Buildings*, 2019, 188-189, 137-148. <https://doi.org/10.1016/j.enbuild.2019.02.012>
- [40] M Schweiker, M Shukuya. Comparison of theoretical and statistical models of air-conditioning-unit usage behaviour in a residential setting under Japanese climatic conditions. *Building and Environment*, 2009, 44(10), 2137-

2149. <https://doi.org/10.1016/j.buildenv.2009.03.004>

[41] Goldsworthy, M. Towards a Residential Air-Conditioner Usage Model for Australia. *Energies* 2017, 10, 1256. <https://doi.org/10.3390/en10091256>

[42] J Tanimoto, A Hagishima. State transition probability for the Markov Model dealing with on/off cooling schedule in dwellings. *Energy and Buildings*, 2005, 37(3), 181-187. <https://doi.org/10.1016/j.enbuild.2004.02.002>

[43] J Wang, J Zhu, Z Ding, P X W Zou, J Li. Typical energy-related behaviors and gender difference for cooling energy consumption. *Journal of Cleaner Production*, 2019, 238, 117846. <https://doi.org/10.1016/j.jclepro.2019.117846>

[44] J Tanimoto, A Hagishima. State transition stochastic model for predicting off to on cooling schedule in dwellings as implemented using a multilayered artificial neural network. *Journal of Building Performance Simulation*, 2012, 5(1), 45-53. <https://doi.org/10.1080/19401493.2010.533388>

[45] Y Zhang, B K Teoh, M Wu, J Chen, L Zhang. Data-driven estimation of building energy consumption and GHG emissions using explainable artificial intelligence. *Energy*, 2023, 262, 125468. <https://doi.org/10.1016/j.energy.2022.125468>

[46] A A Al-Shargabi, A Almhafdy, D M Ibrahim, M Alghieth, F Chiclana. Buildings' energy consumption prediction models based on buildings' characteristics: Research trends, taxonomy, and performance measures. *Journal of Building Engineering*, 2022, 54, 104577. <https://doi.org/10.1016/j.job.2022.104577>

[47] L Lei, W Chen, B Wu, C Chen, W Liu. A building energy consumption prediction model based on rough set theory and deep learning algorithms. *Energy and Buildings*, 2021, 240, 110886. <https://doi.org/10.1016/j.enbuild.2021.110886>

[48] P Anand, C Deb, K Yan, J Yang, D Cheong, C Sekhar. Occupancy-based energy consumption modelling using machine learning algorithms for institutional buildings. *Energy and Buildings*. 2021, 252, 111478. <https://doi.org/10.1016/j.enbuild.2021.111478>

[49] Z Wang, Y Ding. An occupant-based energy consumption prediction model for office equipment. *Energy and Buildings*, 2015, 109, 12-22. <https://doi.org/10.1016/j.enbuild.2015.10.002>

[50] X Li, R Yao. A machine-learning-based approach to predict residential annual space heating and cooling loads considering occupant behaviour. *Energy*, 2020, 212, 118676. <https://doi.org/10.1016/j.energy.2020.118676>

[51] Zhou, X., Mei, Y., Liang, L., Mo, H., Yan, J., & Pan, D. Modeling of occupant energy consumption behavior based on human dynamics theory: A case study of a government office building. *Journal of Building Engineering*, 2022, 58, 104983. <https://doi.org/10.1016/j.job.2022.104983>

[52] Richardson, I., Thomson, M., Infield, D., & Clifford, C. Domestic electricity use: A high-resolution energy demand model. *Energy and Buildings*, 2010, 42(10), 1878-1887. <https://doi.org/10.1016/j.enbuild.2010.05.023>

[53] L Zhang, J Zhang, P Ren, L Ding, W Hao, C An, A Xu. Analysis of energy consumption prediction for office buildings based on GA-BP and BP algorithm. *Case Studies in Thermal Engineering*, 2023, 50, 103445. <https://doi.org/10.1016/j.csite.2023.103445>

[54] L Wang, D Xie, L Zhou, Z Zhang. Application of the hybrid neural network model for energy consumption

prediction of office buildings. *Journal of Building Engineering*, 2023, 72, 106503.

<https://doi.org/10.1016/j.jobe.2023.106503m>

[55] Y Zhang, Y Zhang, Yi Zhang, C Zhang. Effect of physical, environmental, and social factors on prediction of building energy consumption for public buildings based on real-world big data. *Energy*, 2022, 261, 125286.

<https://doi.org/10.1016/j.energy.2022.125286>

[56] Z Dong, J Liu, B Liu, K Li, X Li. Hourly energy consumption prediction of an office building based on ensemble learning and energy consumption pattern classification. *Energy and Buildings*, 2021, 241, 110929.

<https://doi.org/10.1016/j.enbuild.2021.110929>

[57] R Lei, J Yin. Prediction method of energy consumption for high building based on LMBP neural network. *Energy Reports*, 2022, 8, 1236-1248. <https://doi.org/10.1016/j.egyr.2022.02.071>

[58] L Wang, D Xie, L Zhou, Z Zhang. Application of the hybrid neural network model for energy consumption prediction of office buildings. *Journal of Building Engineering*, 2023, 106503.

<https://doi.org/10.1016/j.jobe.2023.106503>

[59] I Amalou, N Mouhni, A Abdali. Multivariate time series prediction by RNN architectures for energy consumption forecasting. *Energy Reports*, 2022, 8, 1084-1091. <https://doi.org/10.1016/j.egyr.2022.07.139>

[60] J Kim, Y Kwak, S Mun, J Huh. Electric energy consumption predictions for residential buildings: Impact of data-driven model and temporal resolution on prediction accuracy. *Journal of Building Engineering*, 2022, 62, 105361. <https://doi.org/10.1016/j.jobe.2022.105361>

[61] C Tarmanini, N Sarma, C Gezezin, O Ozgonenel. Short term load forecasting based on ARIMA and ANN approaches. *Energy Reports*, 2023, 9, 550-557. <https://doi.org/10.1016/j.egyr.2023.01.060>

[62] E Jahani, K Cetin, I H Cho. City-scale single family residential building energy consumption prediction using genetic algorithm-based Numerical Moment Matching technique. *Building and Environment*, 2020, 172, 106667.

<https://doi.org/10.1016/j.buildenv.2020.106667>

[63] Amasyali K, El-Gohary N. Machine learning for occupant-behavior-sensitive cooling energy consumption prediction in office buildings. *Renewable and Sustainable Energy Reviews*, 2021, 142, 110714.

<https://doi.org/10.1016/j.rser.2021.110714>

[64] P Lin, L Zhang, J Zuo. Data-driven prediction of building energy consumption using an adaptive multi-model fusion approach. *Applied Soft Computing*, 2022, 129, 109616. <https://doi.org/10.1016/j.asoc.2022.109616>

[65] X Liang, S Chen, X Zhu, X Jin, Z Du. Domain knowledge decomposition of building energy consumption and a hybrid data-driven model for 24-h ahead predictions. *Applied Energy*, 2023, 344, 121244.

<https://doi.org/10.1016/j.apenergy.2023.121244>

[66] C Tian, C Li, G Zhang, Y Lv. Data driven parallel prediction of building energy consumption using generative adversarial nets. *Energy and Buildings*, 2019, 186, 230-243. <https://doi.org/10.1016/j.enbuild.2019.01.034>

[67] S Norouziasl, A Jafari. Identifying the most influential parameters in predicting lighting energy consumption in office buildings using data-driven method. *Journal of Building Engineering*, 2023, 72, 106590.

<https://doi.org/10.1016/j.jobe.2023.106590>

- [68] T Gao, D Niu, Z Ji, L Sun. Mid-term electricity demand forecasting using improved variational mode decomposition and extreme learning machine optimized by sparrow search algorithm. *Energy*, 2022, 261, 125328. <https://doi.org/10.1016/j.energy.2022.125328>
- [69] G Zhang, Y Ge, X Pan, M Sadat Afsharzadeh, M Ghalandari. Optimization of energy consumption of a green building using PSO-SVM algorithm. *Sustainable Energy Technologies and Assessments*, 2022, 53, 102667. <https://doi.org/10.1016/j.seta.2022.102667>
- [70] Y Xu, N Chen, X Shen, L Xu, Z Pan, F Pan. Proposal and experimental case study on building ventilating fan fault diagnosis based on cuckoo search algorithm optimized extreme learning machine. *Sustainable Energy Technologies and Assessments*, 2021, 45, 100975. <https://doi.org/10.1016/j.seta.2020.100975>
- [71] Y Huang, C Li. Accurate heating, ventilation and air conditioning system load prediction for residential buildings using improved ant colony optimization and wavelet neural network. *Journal of Building Engineering*, 2021, 35, 101972. <https://doi.org/10.1016/j.jobbe.2020.101972>
- [72] M Ghalambaz, R Jalilzadeh Yengejeh, A H Davami. Building energy optimization using Grey Wolf Optimizer (GWO). *Case Studies in Thermal Engineering*, 2021, 27, 101250. <https://doi.org/10.1016/j.csite.2021.101250>
- [73] C Zhang, C Ji, L Hua, H Ma, M S Nazir, T Peng. Evolutionary quantile regression gated recurrent unit network based on variational mode decomposition, improved whale optimization algorithm for probabilistic short-term wind speed prediction. *Renewable Energy*, 2022, 197, 668-682. <https://doi.org/10.1016/j.renene.2022.07.123>
- [74] C Zhang, L Ma, Z Luo, X Han, T Zhao. Forecasting building plug load electricity consumption employing occupant-building interaction input features and bidirectional LSTM with improved swarm intelligent algorithms. *Energy*, 2024, 288, 129651. <https://doi.org/10.1016/j.energy.2023.129651>
- [75] B Doherty, K Trenbath. Device-level plug load disaggregation in a zero energy office building and opportunities for energy savings. *Energy and Buildings*, 2019, 204, 109480. <https://doi.org/10.1016/j.enbuild.2019.109480>
- [76] S Pan, B Yang, S Wang, Z Guo, L Wang, J Liu, S Wu. Oil well production prediction based on CNN-LSTM model with self-attention mechanism. *Energy*, 2023, 284, 128701. <https://doi.org/10.1016/j.energy.2023.128701>
- [77] A Wan, Q Chang, K Al-Bukhaiti, J He. Short-term power load forecasting for combined heat and power using CNN-LSTM enhanced by attention mechanism. *Energy*, 2023, 282, 128274. <https://doi.org/10.1016/j.energy.2023.128274>
- [78] B Cui, M Liu, S Li, Z Jin, Y Zeng, X Lin. Deep learning methods for atmospheric PM2.5 prediction: A comparative study of transformer and CNN-LSTM-attention. *Atmospheric Pollution Research*, 2023, 14(9), 101833. <https://doi.org/10.1016/j.apr.2023.101833>
- [79] W H Chung, Y H Gu, S J Yoo. District heater load forecasting based on machine learning and parallel CNN-LSTM attention. *Energy*, 2022, 246, 123350. <https://doi.org/10.1016/j.energy.2022.123350>
- [80] E M Saoudi, J Jaafari, S J Andaloussi. Advancing human action recognition: A hybrid approach using attention-based LSTM and 3D CNN. *Scientific African*, 2023, 21, e1796. <https://doi.org/10.1016/j.sciaf.2023.e01796>

- [81] D Kapetanakis, E Mangina, D P Finn. Input variable selection for thermal load predictive models of commercial buildings. *Energy and Buildings*, 2017, 137, 13-26. <https://doi.org/10.1016/j.enbuild.2016.12.016>
- [82] J Zhang, Y Huang, H Cheng, H Chen, L Xing, Y He. Ensemble learning-based approach for residential building heating energy prediction and optimization. *Journal of Building Engineering*, 2023, 67, 106051. <https://doi.org/10.1016/j.jobe.2023.106051>
- [83] Z Dong, J Liu, B Liu, K Li, X Li. Hourly energy consumption prediction of an office building based on ensemble learning and energy consumption pattern classification. *Energy and Buildings*, 2021, 241, 110929. <https://doi.org/10.1016/j.enbuild.2021.110929>
- [84] J Ling, B Zhang, N Dai, J Xing. Coupling input feature construction methods and machine learning algorithms for hourly secondary supply temperature prediction. *Energy*, 2023, 278, 127459. <https://doi.org/10.1016/j.energy.2023.127459>
- [85] J Ling, N Dai, J Xing, H Tong. An improved input variable selection method of the data-driven model for building heating load prediction. *Journal of Building Engineering*, 2021, 44, 103255. <https://doi.org/10.1016/j.jobe.2021.103255>