Energy & Buildings 240 (2021) 110886

Contents lists available at ScienceDirect

Energy & Buildings

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

A building energy consumption prediction model based on rough set theory and deep learning algorithms

Lei Lei [a], Wei Chen [a], Bing Wu [a], Chao Chen [a], Wei Liu [b,*]

[a] School of Architecture and Transportation Engineering, Guilin University of Electronic Technology, 1 Jinji Road, Guilin 541004, China
[b] Division of Sustainable Buildings, Department of Civil and Architectural Engineering, KTH Royal Institute of Technology, Brinellvägen 23, Stockholm 100 44, Sweden

ARTICLE INFO

Article history:
Received 3 October 2020
Revised 11 January 2021
Accepted 3 March 2021
Available online 10 March 2021

Keywords:
Building energy consumption
Prediction model
Rough set
Deep learning
Attribute reduction

ABSTRACT

The efficient and accurate prediction of building energy consumption can improve the management of power systems. In this paper, the rough set theory was used to reduce the redundant influencing factors of building energy consumption and find the critical factors of building energy consumption. These key factors were then used as the input of a deep neural network with a "deep" architecture and powerful capabilities in extracting features. Building energy consumption is output of the deep neural network. This study collected data from 100 civil public buildings for rough set reduction, and then collected data from a laboratory building of a university in Dalian for nearly a year to train and test deep neural networks. The test included both the short-term and medium-term predictions of building energy consumption. The prediction results of the deep neural network were compared with that of the back propagation neural network, Elman neural network and fuzzy neural network. The results show that the integrated rough set and deep neural network was the most accurate. The method proposed in this study could provide a practical and accurate solution for building energy consumption prediction.

1. Introduction

With the rapid development of the global economy, building energy consumption keeps increasing and it now accounts for more than one-third of total global energy consumption [1]. Therefore, reducing building energy consumption to relieve energy pressure has become particularly important for realizing sustainable goals such as affordable energy and sustainable cities. Accurate prediction of building energy consumption is one of the effective means to reduce building energy consumption, since it helps achieve better control of power system and improve energy utilization [2].

The methods of performing building energy consumption prediction can be generally divided into physical-model based methods and data-driven methods. Among them, physical-model based methods are also called white-box based approaches, which use the principles of physics to evaluate the building energy consumption. For example, to calculate the energy consumption of the heating ventilation, and air conditioning (HVAC) system, the general methods are variable base degree day methods, equivalent full load hours method and bin method. The commonly used software for building energy simulation includes EnergyPlus, equest, and octet, etc.[3] In simulating the building energy consumption, the physical-model based methods require detailed input information, such as physical features of the building, HVAC systems, equipment, and occupants' schedule, etc.[4]. The informative input would help achieve accurate prediction of building energy consumption, but as well makes the modeling process quite complicated and slow.

Data-driven methods are divided into two categories: statistical methods and machine learning methods. The statistical methods are also called grey box models, which are generally in the form of regression analysis prediction [5] or time series-based prediction. Regression analysis simulates and predicts energy consumption by establishing mathematical expressions between independent and dependent variables. Guo et al. [6] proposed an improved multivariable linear regression model to predict the daily mean cooling load of office buildings in which three main measures, including the principal component analysis of meteorological factors, cumulative effect of high temperature and dynamic two-step correction. Pombeiro et al. [7] compared and evaluated the prediction results of several low-complexity models for the energy consumption of an institutional building. Their research results indicated that both the developed fuzzy systems and neural networks have better performance and accuracy indicators than linear regression models. Time series-based analysis is a statistical

* Corresponding author.
E-mail address: weiliu2@kth.se (W. Liu).

https://doi.org/10.1016/j.enbuild.2021.110886
0378-7788/© 2021 The Author(s). Published by Elsevier B.V.

Nomenclature

U	Domain
A	Non-empty finite set of attributes
K	The range of all attributes
f	Information function
C	Conditional attribute
D	Decision attribute
V	visible layer of restricted Boltzmann machine
H	hidden layer of restricted Boltzmann machine
$W_{m\times n}$	the weight matrix between the visible layer and the hidden layer
b	the offset of the visible node
c	the offset of the hidden node

Acronyms

HVAC	heating, ventilating, and air conditioning
SVM	support vector machine
ANN	artificial neural network
BP	back propagation
RS	rough set
DBN	deep belief neural network
RBM	restricted Boltzmann machine
FNN	fuzzy neural network
RS-BP	rough set combined with BP neural network
RS-SRN	rough set combined with Elman neural network
RS-FNN	rough set combined with fuzzy neural network
RS-DBN	rough set combined with deep belief neural network

method for dynamic data processing. Jallal et al. [8] proposed a time series forecasting algorithm based on a hybrid neuro-fuzzy inference system to predict building energy consumption. The results show that this method is more robust than the traditional adaptive neuro-fuzzy inference system models. However, the prediction process of this method is more complicated, and the efficiency of calculating the energy consumption is not high. In general, statistical methods have the advantage of a simple structure and relatively easy model establishment. However, at the same time, due to the complex interaction between input elements in the gray-box model [9], problems such as a low calculation efficiency and low prediction accuracy are likely to occur.

To overcome the deficiencies of the physical-model based methods and the statistical methods, machine learning method has been developed [10]. Machine learning methods for building energy consumption predictions generally include support vector machines (SVMs), gray system theory, and artificial neural networks (ANNs):

- SVMs is a learning method based on structural risk minimization criteria. It can predict building energy consumption with limited samples. Zhong et al. [1] proposed a support vector regression method that improves the accuracy and generalization ability of building energy consumption prediction by looking for the optimal feature space. The model only processes a small sample of building energy consumption data for one month, and it is only suitable for short-term predictions of building energy consumption. Dong et al. [11] used an SVM method to predict the monthly building energy consumption of four commercial buildings in the tropics. The error of the prediction results is small, which is related to the small data pool used in the study. The whole data profile in this study consisting of only 4 years' monthly utility bills may include less scattered or abnormal data. The SVM method is suitable for nonlinear high-dimensional pattern recognition, but this method is memory and computational demanding, which make it difficult to implement for large-scale training samples.
- Gray system theory predicts the development of building energy consumption by analyzing the correlation between system factors to generate a data sequence with strong regularity. In terms of the application of gray system theory, Robert and Uwe [12] used gray system theory to analyze the thermal process of buildings and predict the thermal energy consumption of HVAC system. Guo et al. [13] proposed a method to predict the energy consumption of domestic heat pump water heaters based on the gray system theory. It should be noted that although the gray system theory is simple to calculate, it has a narrow scope of application. It is rarely used to directly predict the energy consumption of the entire building and is generally only suitable for short-term prediction of the energy consumption of part of the building system.
- ANNs are more popular in building energy consumption prediction research because of their strong self-learning and good nonlinear function fitting ability, which can maintain a relatively high accuracy in energy consumption prediction. Wei et al. [14] combined blind system identification and feed-forward neural network to predict the occupancy level and energy consumption of office buildings. Ekici and Aksoy [15] used back propagation (BP) neural network to predict the heating energy requirements of three different buildings. Rahman et al. [16] proposed a recurrent neural network model that can predict the power consumption of commercial and residential buildings for periods longer than a week at one-hour resolution intervals. Li et al. [17] combined a genetic algorithm and an adaptive neural fuzzy inference system to predict the energy consumption of campus buildings. These researches on building energy consumption prediction based on shallow neural networks show that neural networks have better prediction accuracy than other methods. But they also show that shallow neural networks are prone to losing the original data information, and the selection of its network topology is generally determined by trial and error, which is difficult to find the optimal network topology.

In addition to the three main machine learning algorithms mentioned above, some studies have applied decision trees [18,19], extreme learning machines [20,21] and other artificial intelligence algorithms to building energy consumption prediction. In general, the existing machine learning algorithms are more advanced than physical models and statistical methods in all aspects. However, the existing machine learning algorithms have the problems of inaccurate prediction results, low calculation efficiency and limited application range. With the continuous development of the artificial intelligence, there is now a deep learning algorithm, which is expected to make up for the shortcomings of existing building energy consumption prediction methods and bring accurate predictions to building energy consumption.

Building energy consumption is affected by many factors, such as weather conditions, building thermal environment, and building energy usage. The data of building energy consumption have a high degree of uncertainty and randomness, which makes it difficult for the prediction model to be of low bias and variance. Deep learning is able to make full use of the training data to extract the building energy consumption features, which further leads to accurate prediction of building energy consumption. Compared with the shallow neural network, the deep neural network contains multiple hidden layers and has a "deep" architecture. These features make it a very popular artificial intelligence algorithm at present. In medicine [22,23], finance [24,25], transportation [26], environment [27] and many other fields, there is related research and applications of deep learning. Using deep learning as a post information recognition system for predictive models, the features of energy consumption data are extracted through the powerful feature extraction capabilities and "depth" architecture of deep learning. In this way, high-precision building energy consumption prediction results can be obtained.

2. Research methodology

The prediction of building energy consumption can be divided into short-term prediction, medium-term prediction and long-term prediction according to the length of the prediction time scale [28]. The forecast range of general short-term forecasts is several hours to one week [29,30], the forecast range of medium-term forecasts is from several weeks to several months [31], and the realization of long-term forecasts is usually based on the establishment of annual forecasts [32]. Due to data limitation, this study focused on the short-term and mid-term forecasts.

Fig. 1 shows the overall research framework. The parameters C_i (i = 1, 2, ..., 20) are the 20 inputs, which represents the hourly based weather (air temperature, humidity, wind speed, and solar radiation), building information (levels, area, orientation, window-to-wall ratio, thermal conductivity of the external wall and roof, shading coefficient, and length to width ratio), building utilization (lighting fixtures, occupant density, and indoor air temperature), and air conditioning system (fresh air flowrate, COP, air supply temperature, fan efficiency, and pump efficiency). Since there is larger amount of input data, this investigation firstly applied rough set theory, which was proposed by Pawlak [33], to eliminate the redundant data and identify the effective hourly input data Z_i (i = 1, 2, ..., m; m $\leq$ 20) that mainly determines the building energy consumption. For the short-term, this study also developed the neural network without rough sets theory as a comparison. For the mid-term prediction, the hourly data Z_i are processed into daily data T_i (i = 1, 2, ..., m) for the neural network. In this study, the rough set theory is used as the front-end processor of the prediction model, and deep learning is used as the post-information recognition system of the prediction model, which appropriately combines the advantages of both. In this way, not only simplifies the sample structure and improves the generalization ability and running speed of the network, but also strengthens the anti-interference ability and prediction accuracy of the model. This study collected data from 100 civil public buildings for rough set reduction, and then collected data from a laboratory building of a university in Dalian for nearly a year to train and test deep neural networks. The following sections introduces the rough set theory, deep learning, and the prediction process in sequence.

2.1. Rough set theory

The rough set theory is mainly used for learning and induction of incomplete data and uncertain knowledge [34]. It is good at mining the potential relationship between knowledge and finding the key attributes, which make it very suitable for extracting the important influencing factors of building energy consumption.

This study defined $S = (U, A, K, f)$ as a knowledge expression system to classify objects. U = {x_1, x_2, x_3.....x_n} is the domain, where x_i is an object that represents a set of data samples. A = {a_1, a_2, a_3, ...a_n} is a nonempty finite set of attributes and $a_i(x_j)$ is the value of $\times$ recorded on attribute a. The elements in set A refer to the input parameters and building energy consumption. $a(x)$ are the data corresponding to each impact factor. K is the value range of attribute set A. f is the information function set of U and A, and represents a mapping of $U \times A \rightarrow K$, which assigns an information value to each attribute of each object, namely, $\forall a \in A, x \in U$, $f(x, a) \in K$.

If the nonempty finite set of attributes A is composed of a conditional attribute set C and a decision attribute D and satisfies $A = C \cup D, C \cap D \neq \varnothing$, then S is called a decision system. Generally, dealing with decision problems by the rough set method can be expressed in the form of decision table. In this study, the decision table of the information system was expressed as S= (U, C, D, K, f). Each row in the decision table describes an object (a set of data samples, which is called a decision rule), and each column represents an attribute of the object (conditional attribute set C or decision attribute D). In this study, the conditional attribute set C represents various factors that affect the building energy consumption, and the decision attribute D represents building energy consumption.

In engineering applications, the range of the conditional attributes and decision attributes of the problem mostly involve continuous values. However, when the rough set method is used to deal with the decision table, the attribute values in the decision table must be discrete. Therefore, the data need to be discretized for establishing the decision table. This study discretizes the data by dividing the range into a finite number of regions by selecting breakpoints and each region corresponds to an attribute value. For example, x_l (l = 1, 2, ..., n-1) is selected as a breakpoint in the range of values, and then, the n-1 breakpoint divides the continuous range into n attribute values, which can reduce the complexity of the problem. The discretization of continuous attributes can be divided according to expert experience or automatically divided by the system according to certain rules. The most commonly used discretization methods for continuous attribute are equal distance division, equal frequency division, the L-method, the P-method, the Naivescaler algorithm, the SemiNaivescaler algorithm, and Boolean reasoning [35]. This study adopted the equidistant partitioning to discretize the continuous attributes in the attribute reduction algorithm of rough sets.

In the information system of the raw rough set decision table, some of the conditional attributes are critical to the information system of the decision table, while others may be redundant. The purpose of attribute reduction is to simplify the information system by finding the necessary conditional attributes and eliminating the redundant and unimportant conditional attributes in the information system. The attribute reduction of a decision table is not unique, there are many minimum relative reduction results. Therefore, to ensure that the classification ability of the information system after attribute reduction remains unchanged, multiple minimum relative reduction results was considered.

2.2. Deep learning

Deep learning is a branch of machine learning. The basic structure of deep learning is deep neural networks. The difference between deep neural networks and shallow neural networks is that shallow neural networks generally have only two to three layers of neural networks. Therefore, shallow neural network has limited ability to express complex functions. However, deep learning has five or more layers of neural networks and introduces more effective algorithms [36] that can further improve the accuracy of building energy consumption prediction. In this study, the deep

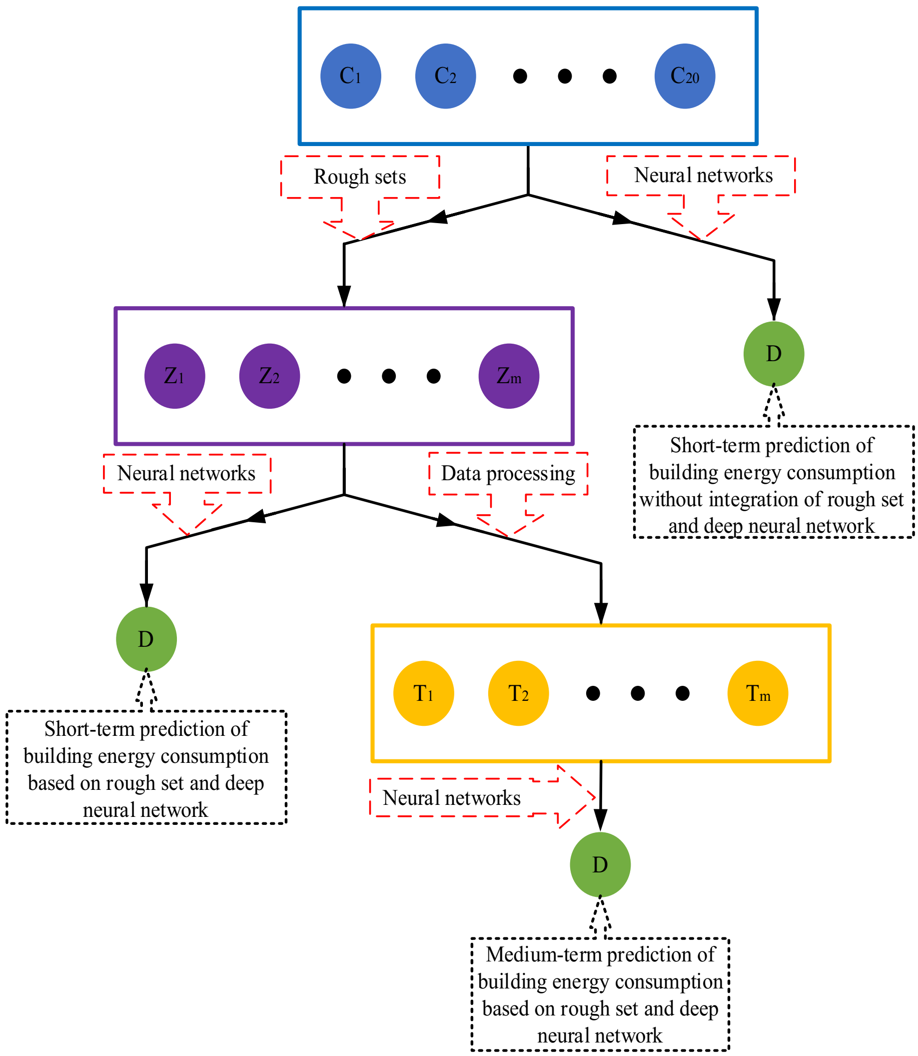

Fig. 1. Research outline.

belief neural network (DBN), which is commonly used in deep learning, is used to predict building energy consumption.

2.2.1. Structure of the DBN

The DBN is made up of multiple restricted Boltzmann machine (RBM) stacks. The structure of the RBM can be represented by the bipartite graph including the visible layer and the hidden layer, as shown in Fig. 2. There are no links between the nodes within each layer. Layer V represents the input data that is visible and layer H is a hidden layer. In RBM, nodes are random binary variable nodes (with values of only 0 or 1), and the full probability distribution $p(V, H)$ satisfies the Boltzmann distribution. Different from the traditional feed-forward neural network, the link between the visible layer and the hidden layer of the RBM is uncertain and fully linked. In other words, the values between layers can carry out the two-way propagation from the hidden layer to the visible layer and vice versa.

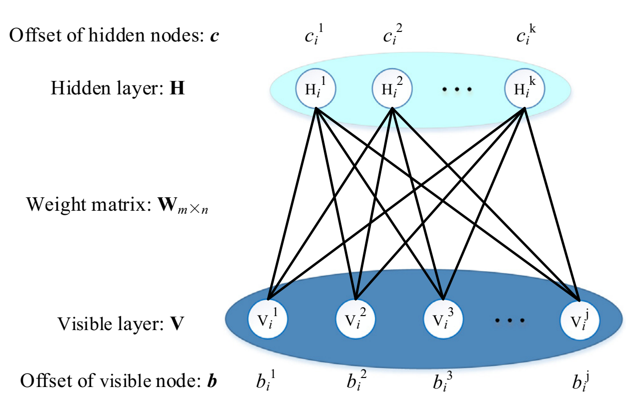

Fig. 2. Structure of RBM model.

The above RBM structure has n visible nodes and m hidden nodes. Each visible node is only related to m hidden nodes. The nodes in the same layer are independent. This feature makes the training process of RBM easier. RBM has several parameters, which are weight matrix $\boldsymbol{W}_{m\times n}$, the offset of the visible node $b_i = (b_i^1, b_i^2, \ldots, b_i^j)$, and the offset of the hidden node $c_i = (c_i^1, c_i^2, \ldots, c_i^k)$. These parameters determine how the RBM network encodes an n-dimensional input into a m-dimensional output.

RBMs are stacked like bricks to build a network, and this network is called a DBN. A typical DBN network is shown in Fig. 3. These RBMs in the DBN network have only one visible layer and one hidden layer, and while the layers are connected to each other. The $\boldsymbol{W}$, $\mathbf{b}$, and $\mathbf{c}$ of each hidden layer will be trained to represent high-level features.

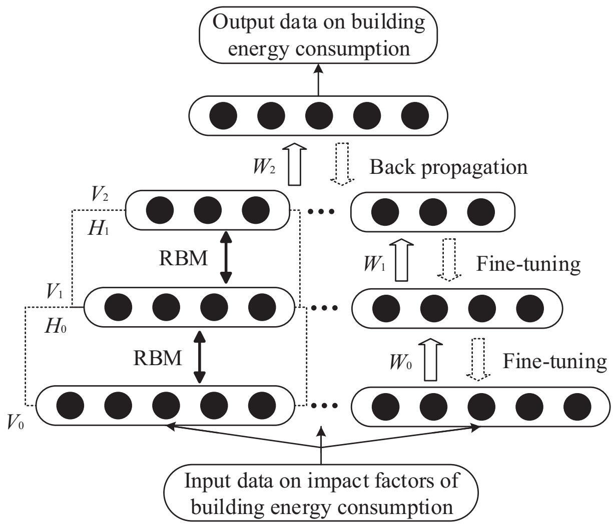

Fig. 3. Structure and training process of DBN.

2.2.2. Training process of DBN

The training process of an RBM is to determine the probability distribution with the highest probability of generating training samples. Since the decisive factor for the probability distribution is the weight *W*, the goal of training an RBM is to find the best weight. This investigation adopted the greedy layer-wise training method to train the RBMs one by one.

The training process of the DBN is mainly divided into two parts: pretraining and fine-tuning. In the first part, unsupervised and spontaneous training is performed on each RBM to ensure that feature information is retained as much as possible when feature vectors are mapped to different feature spaces. The input layer of the first RBM is the input layer of the entire neural network. The hidden layer H_i can be regarded as the visible layer of H_{i+1}, which is denoted as V_{i+1}. The pretraining of the greedy layer-wise training method involves the following steps.

- Step 1: Train the first RBM for all training samples.
- Step 2: Fix the weight and offset of the first RBM, and then train the second RBM with $V_1 = H_0$.

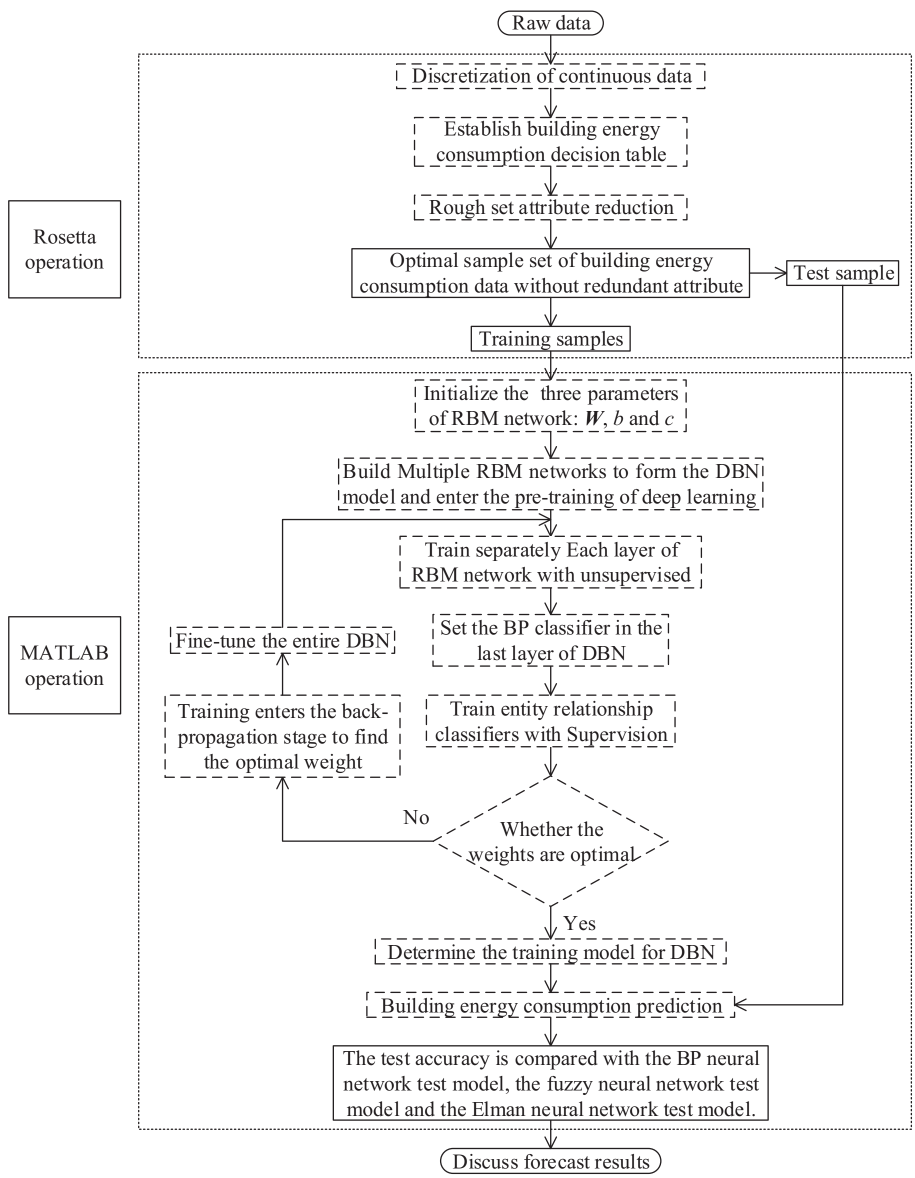

Fig. 4. Workflow of prediction model for building energy consumption based on rough set and deep learning.

Table 1
Raw data used for attribute reduction in rough sets.

Sample	Outdoor temperature	Relative humidity	Wind speed	Solar radiation	Number of floors	Building area	Building orientation	Window-to-wall area ratio	Heat transfer coefficient of external walls	Shading coefficient	Building length-to-width ratio	Heat transfer coefficient of roof	Lighting power density	Personnel density	Indoor temperature	Fresh air volume per capita	COP of the chiller	Supply air temperature	Fan efficiency	Pump efficiency	Energy consumption
	°C	%	m/s	W/m^2		m^2	degree	%	W/(m^2·K)			W/(m^2·K)	W/m^2	person / m^2	°C	m^3/h. person		°C	%	%	kWh
1	25.59	86.8	0.83	155.56	4	3520	234	16.3	0.84	0.68	1.51	0.56	7.86	0.25	24.18	28	6.31	20.4	0.66	0.75	39.6
2	25.49	80.1	0.56	310.2	5	5630	263	18.2	0.65	0.59	1.65	0.74	4.69	0.65	24.75	29	6.75	21	0.56	0.56	36.8
3	25.25	79.6	1.20	382.4	4	4856	256	26.8	0.87	0.68	1.85	0.86	11.65	0.99	23.33	36	6.4	21.3	0.59	0.52	88
4	25.05	79.7	0.42	410.84	6	4586	281	17.5	0.96	0.49	1.64	0.94	13	1.11	24.78	25	6.17	21.7	0.85	0.59	67.6
5	24.67	82.1	0.85	390.41	5	7854	256	20.3	0.70	0.95	1.95	0.65	15.92	1.10	23.28	25	5.87	21.9	0.89	0.8	58
6	24.48	80.9	0.80	329.08	7	8561	269	21.2	1.36	0.87	2.64	0.74	6.94	0.48	23.69	28	5.61	22.1	0.90	0.56	68
7	24.98	80.2	0.59	228.42	6	8463	249	18.9	1.29	0.56	2.59	0.60	6.94	0.69	24.21	23	5.61	22.3	0.57	0.59	37.6
8	29.1	80.5	0.34	102.78	5	8564	268	14.6	1.45	0.95	3.15	0.98	8.95	0.59	26.88	37	5.47	22.5	0.67	0.5	70.8
9	29.37	81.1	0.23	266.66	6	8856	235	15	0.71	0.68	1.65	1.00	12.6	0.30	28.98	26	5.47	22.5	0.58	0.68	68.8
10	29.58	86.6	0.62	368.47	4	4685	268	17.5	0.85	0.84	2.98	0.94	12.59	0.54	28.62	25	5.54	17.6	0.59	0.8	97.6
11	30.41	90.1	0.38	403.4	6	3698	264	18.8	0.64	0.50	2.36	0.56	12.96	0.98	29.43	23	5.38	22.8	0.81	0.81	65.2
12	31.22	92.3	0.93	394.32	7	8463	284	17.9	0.69	0.59	1.78	0.59	11.59	0.85	29.69	31	5.29	22.8	0.86	0.82	66.4
13	32.06	94.5	0.57	341.07	5	6584	256	20.1	0.84	0.54	3.01	0.83	9.65	0.76	30.10	27	5.23	22.9	0.79	0.76	73.2
14	32.51	95.1	1.04	246.86	6	7569	248	20.3	0.91	0.68	1.99	0.97	9.05	0.29	30.57	25	5.11	19	0.90	0.79	65.2
15	32.93	97.3	0.93	118.82	4	8465	268	23.4	1.25	0.79	2.64	0.76	9.06	1.02	30.22	23	5.18	18.1	0.81	0.72	96
16	33.52	95.9	0.77	309.94	6	4869	269	16.8	1.24	0.76	3.16	1.00	9.00	1.13	31.29	29	5.2	17.2	0.55	0.61	95.6
17	32.83	96	0.38	277.78	7	5896	281	19.5	1.34	0.71	2.46	0.99	5.92	0.95	30.77	28	5.14	17.8	0.59	0.69	86.8
18	31.1	97.4	1.16	361.12	4	5963	264	17.5	0.94	0.81	2.59	0.94	5.62	0.98	28.05	26	5.14	20.4	0.55	0.59	103.6
19	30.01	96.2	0.47	555.56	5	3874	256	18.8	1.35	0.94	2.76	0.96	6.95	1.01	29.15	23	4.95	20.6	0.88	0.54	121.8
...										...											...
100	29.87	91.8	0.92	538.88	6	7658	265	17.6	0.62	0.93	2.54	0.98	8.32	1.09	26.84	25	5.1	19.4	0.81	0.5	98.4

- Step 3: Repeat the previous two steps for all RBMs.

In the second part, the fine-tuning set the classifier model for the last layer of the DBN. In this study, the classifier model is set to the commonly used BP network. By receiving the output feature vector of the RBM as its input feature vector, the entity relationship classifier is trained with supervision. Each layer of the RBM network can only ensure that the weights in its own layer are optimized for the feature vector mapping of the layer but not for feature vector mapping of the entire DBN. Therefore, the reverse propagation network also propagates the error information from top to bottom of each layer of the RBM to fine-tune the DBN network. The pretraining process of the RBM can be regarded as the initialization the of weight parameters of a deep BP network, which facilitates the DBN to overcome the problem of the BP network's tendency to fall into a local optimum and the long training time due to the random initialization of weight parameters.

2.3. Prediction process of the model

The main workflow of the building energy consumption prediction is shown in Fig. 4. This study firstly determined the discrete breakpoints of each input and discretized the continuous raw data. Through attribute reduction by rough set, the important inputs for building energy consumption was obtained. Then the data sample set was divided into training data and test data. The training data was used for DBN model training, and the test data was used for verifying the model accuracy. The training of the DBN model started by initializing the three parameters $\boldsymbol{W}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ of the RBM. This investigation then performed unsupervised pre-training and unsupervised fine-tuning of the DBN. In this way, the optimal weights for mapping the entire DBN were obtained.

In this study, Rosetta [37], a rough set software, is used to reduce the attributes in the rough set decision table formed by the sample data. Rosetta is a tabular logical data tool based on the rough set theory framework. It is often used for the algorithmic acquisition of the rules and attribute reduction in rough set theory. The attribute reduction method based on the genetic algorithm provided in the software is selected to reduce the 20 inputs of building energy consumption. The remaining conditional attributes after reduction are the corresponding important inputs of building energy consumption. The establishment, training, and test of the neural networks were performed in MATLAB [38].

3. Data collection

The field measurements of data had two parts. The first part was the data collected of 100 civil public buildings for reducing the inputs of building energy consumption (attributes) by rough sets. The second part was the data collected of a laboratory building for training and testing the deep neural network.

3.1. Data collection for attribute reduction in rough sets

In this study, civil public buildings were used as experimental objects. Civil public buildings refer to non-residential buildings for people to carry out various public activities, generally including office buildings, commercial buildings, tourist buildings, science, education, culture and health buildings, and transportation buildings. A total of 20 inputs C_i (i = 1, 2, ..., 20) that represent building energy impact factors and the building energy consumption were collected. For example, the outdoor temperature, relative humidity and wind speed were measured by hot-wire anemometer TES-1341. The solar radiation was measured by solar radiation sensor HOBO S-LIB-M003. The solar radiation refers to the solar radiation outside the building in the natural state. Its value is the total solar radiation intensity, including the direct solar radiation intensity and the solar scattered radiation intensity. The occupancy was estimated by measuring the concentration of CO2 in the building with a carbon dioxide detector Telaire-7001. The supply air temperature of HVAC units was measured by air temperature and humidity sensor HOBO UX100-003. The building energy consumption was obtained by real-time monitoring. Building energy consumption was determined by electricity consumption, which considered the energy used by the HVAC system, lighting fixtures, and electrical appliances.

The data collection were conducted on 100 civilian public buildings in Dalian during the period from May 2017 to September 2017, and one set of test data on building energy consumption and the influencing factors of building energy consumption were collected for each building. The test period for each set of data was one hour. Some of the data are shown in Table 1. Among them, the indoor temperature refers to the average of measured air temperature in the room over one hour. The heat transfer coefficient of the exterior wall is the average heat transfer coefficient of all exterior walls. The shading coefficient refers to the external shading coefficient, which is the ratio of the solar radiation heat gained through an external window with an external shading structure

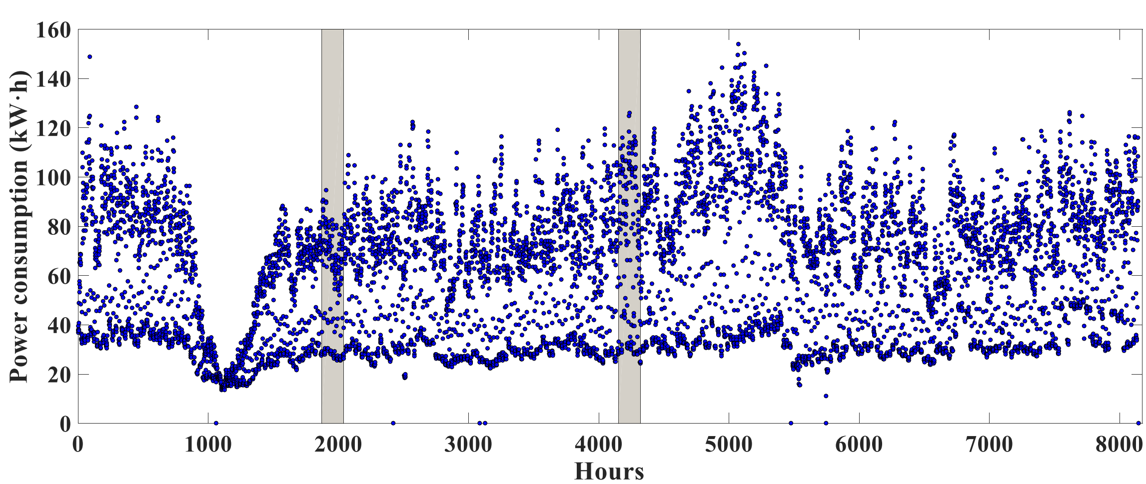

Fig. 5. Hourly energy consumption of the laboratory building.

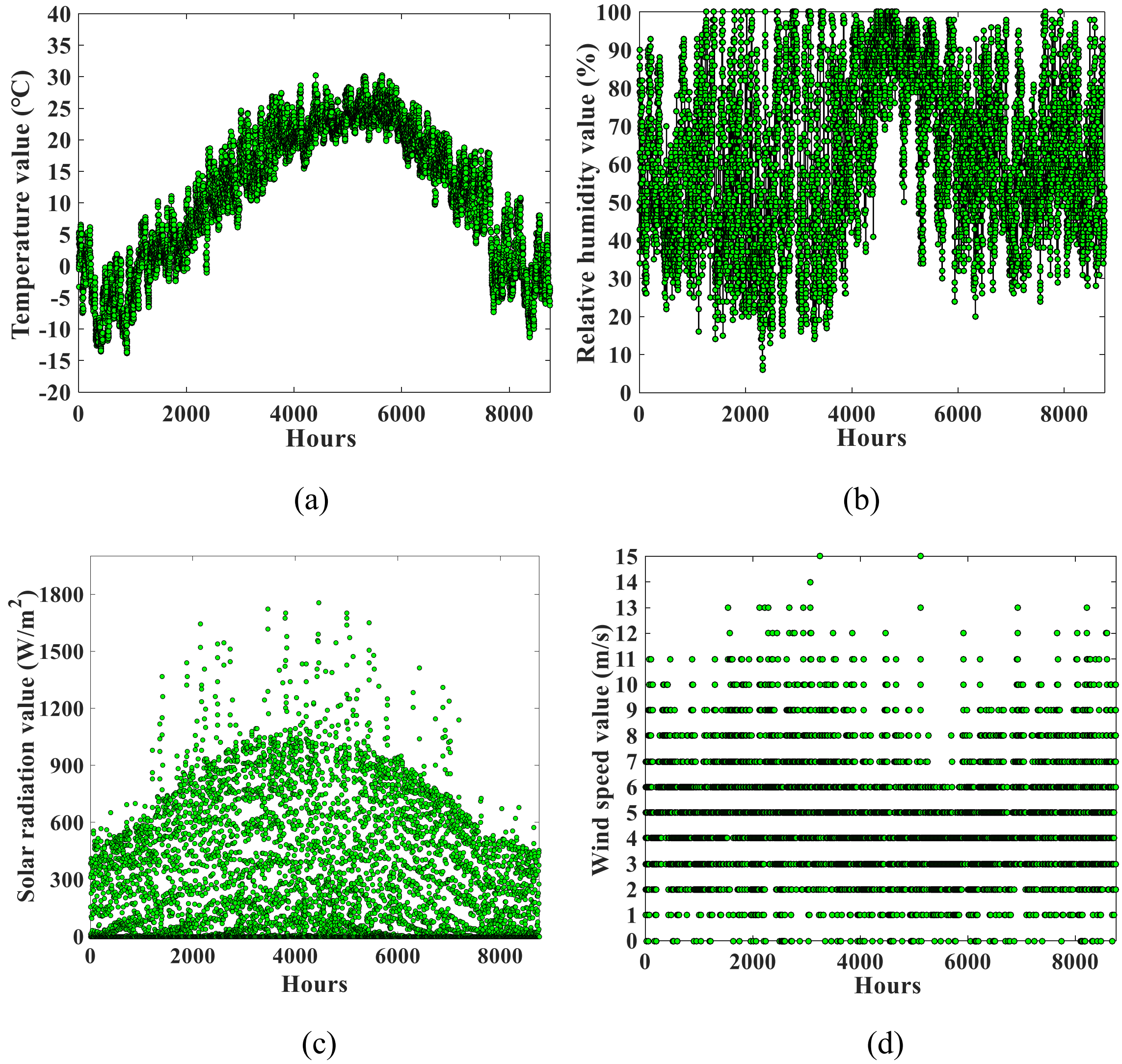

Fig. 6. Hourly meteorological data of Dalian throughout the year (a) air temperature; (b) relative humidity; (c) solar radiation; (d) wind speed.

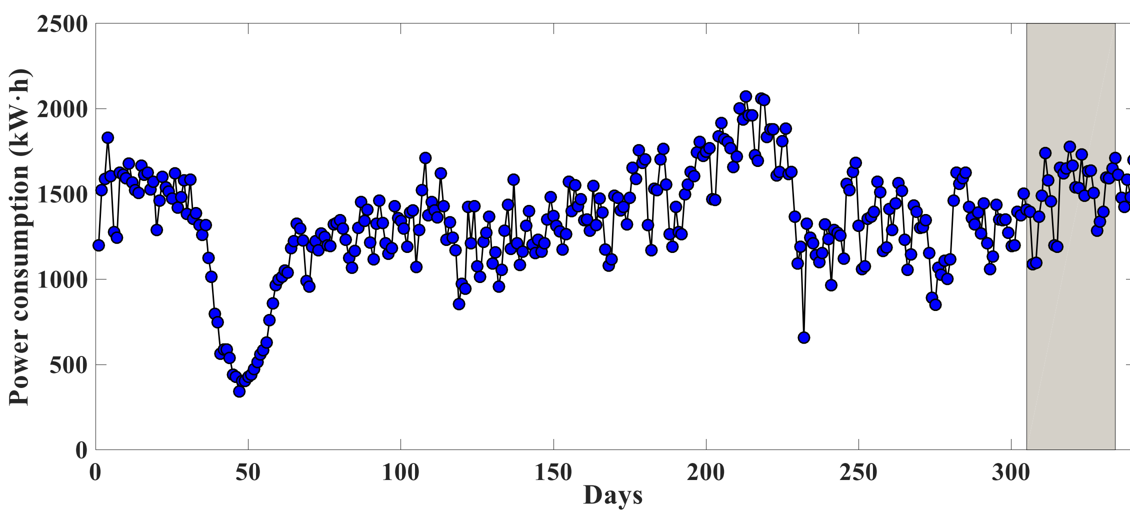

Fig. 7. Daily energy consumption of the laboratory building.

to that through the same external window without an external shading structure. This study adopted a simplified formula and took the shading structure size ratio × as a parameter to calculate the external shading coefficient by SD = $ax^2 + bx + 1$, where a and b

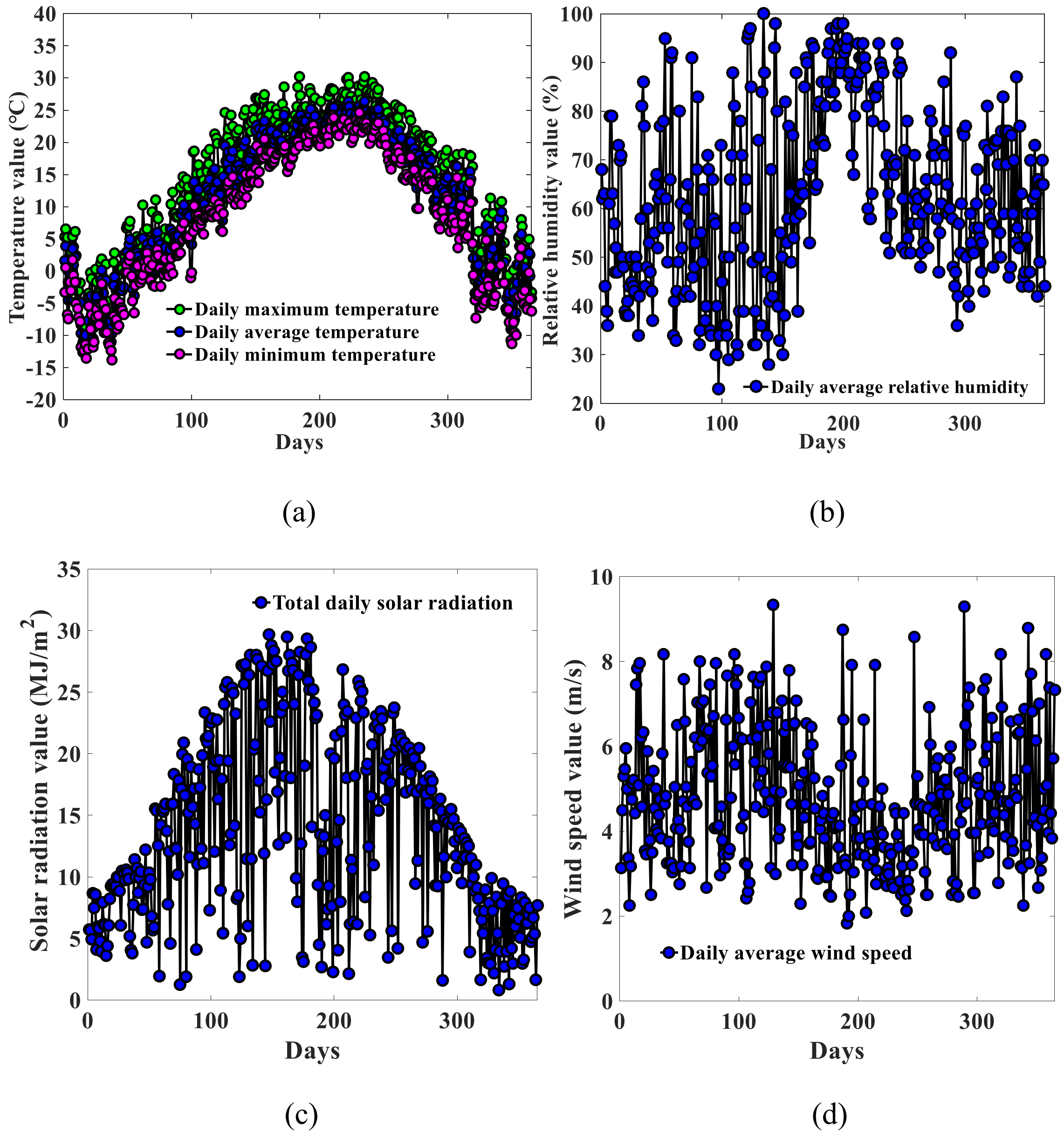

Fig. 8. Daily meteorological data of Dalian throughout the year (a) Daily outdoor average temperature; (b) Daily average humidity; (c) Daily solar radiation; (d) Daily average wind speed.

are regression coefficients [39]. The values of a and b are dependent on the climate zone where the building is located, the basic type of external shading, and the orientation of the windows.

3.2. Data collection for the deep neural network

After reducing the number of inputs by rough sets, the effective inputs Z_i (i = 1, 2, ..., m; m $\leq$ 20) that determines the building energy consumption are obtained. This investigation collected data including the effective inputs Z_i and building energy consumption in a laboratory building of a university in Dalian. The laboratory building is a frame-shear structure, with five floors above ground and one floor underground, with a total construction area of 9533 m^2 and an air-conditioning area of 8579.7 m^2.

Fig. 5 shows 8176 h of hourly building energy consumption of the laboratory building from January 1, 2018 to December 7, 2018. The dots lie in the two shaded areas represent the two periods (168 h each) for testing the short-term predictions. The period from the 1000th hour to the 2000th hour was spring, and the laboratory building had low energy consumption. The season around the 4000th to 5000th hour was summer, which is the period that the laboratory building consumes the most energy throughout the year. Please note that the energy consumption refers to the electrical energy consumption, the energy consumption by the central heating system in winter was no considered. Fig. 6 shows the corresponding hourly meteorological data of Dalian, which is part of the input data. Among them, the period around 4000th to 5000th hour was summer, so the air temperature, relative humidity and solar radiation were relatively high.

The measured hourly data throughout the year is for short-term forecast of building energy consumption. The hourly data were further processed into daily data for the medium-term forecast of building energy consumption. For example, the summation of the hourly energy consumption in a day is the daily energy consumption and average of the hourly outdoor air temperature in a day is the daily outdoor air temperature. Fig. 7 shows the daily energy

Table 2
Discrete breakpoints of each attribute.

Inputs (influence factors of building energy consumption)	Label	Minor impact (1)	Slight impact (2)	Moderate impact (3)	Significant impact (4)
Outdoor temperature	C_1	<=26	(26,29)	(29,32)	>32
Relative humidity	C_2	<=60	(60,75]	(75,90]	>90
Wind speed	C_3	>1.6	(1.6,1.3]	(1.3,1.0]	<=1
Outdoor solar radiation illuminance	C_4	<=150	(150,300]	(300,450]	>450
Number of floors	C_5	<=3	(3,5]	(5,7]	>7
Building area	C_6	<=3000	(3000,6000]	(6000,9000]	>9000
Building orientation (general orientation: south)	C_7	<=240	(240,270]	(270,300]	>300
Window-to-wall area ratio	C_8	<=0.15	(0.15,0.2]	(0.2,0.25]	>0.25
Heat transfer coefficient of external walls	C_9	<=0.8	(0.8,1.1]	(1.1,1.4]	>1.4
Shading coefficient	C_{10}	<=0.5	(0.5,0.65]	(0.65,0.8]	>0.8
Building length-to-width ratio	C_{11}	<=1.5	(1.5,2.5]	(2.5,3.5]	>3
Heat transfer coefficient of roof	C_{12}	<=0.6	(0.6,0.75]	(0.75,0.9]	>0.9
Lighting power density	C_{13}	<=6	(6,9]	(9,12]	>12
Personnel density	C_{14}	<=0.4	(0.4,0.7]	(0.7,1.0]	>1.0
Indoor temperature	C_{15}	<=25	(25,29]	(29,32]	>32
Fresh air volume per capita	C_{16}	<=30	(30,40]	(40,50]	>50
COP of the chiller	C_{17}	>6	(5,6]	(4,5]	(0.7,1.0]
Supply air temperature	C_{18}	>24	(21,24]	(18,21]	(29,32]
Fan efficiency	C_{19}	>0.8	(0.7,0.8]	(0.6,0.7]	<=0.6
Pump efficiency	C_{20}	>0.8	(0.6,0.8]	(0.4,0.6]	<=0.4

Table 3
The levels of energy consumption.

Rank sequence	Characterization state	Energy consumption description
1	Low energy consumption	<=50
2	Moderate energy consumption	(50,80]
3	High energy consumption	(80,110]
4	Ultra-high energy consumption	>110

consumption of 341 days. The shaded part represents the 30-day test period for mid-term prediction of building energy consumption. Fig. 8. Shows the daily meteorological data of Dalian throughout the year.

4. Results

In order to identify the effective inputs for the building energy consumption, this study first applied the rough set with the data collected in section 3.1. The determined effective inputs are used as the inputs of the deep neural network and the building energy consumption is the output of the deep neural network. The data collected in section 3.2 was used for training and testing the deep neural network.

4.1. Identification of effective inputs for building energy consumption

Since rough sets cannot process continuous data, the original data used for attribute reduction in the rough sets must be discretized first. The equidistance division is used as the ideological basis of the discretization process. Combined with the basic characteristics of the data and the degree of centralization, the discrete points of the individual influencing factors of building energy consumption are slightly adjusted. The influencing factors of building energy consumption are used as energy consumption indicators. The value range of each attribute is discretized into four intervals, coded as 1, 2, 3, and 4, which indicate that the input has minor, slight, moderate, and significant impact on the energy consumption, respectively. C_1, C_2, C_3, ..., C_{20} represent the 20 influencing

Table 4
Decision table of building energy consumption in the laboratory building.

Sample	Conditional attributes (inputs) C																				Decision attribute (output) D
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}	C_{17}	C_{18}	C_{19}	C_{20}	Decision value
1	1	3	4	2	2	2	1	2	2	3	2	1	2	1	1	1	1	3	3	2	1
2	1	3	4	3	2	2	2	2	1	2	2	2	1	2	1	1	1	3	4	3	1
3	1	3	3	3	2	2	2	4	2	3	2	3	3	3	1	2	1	2	4	3	3
4	1	3	4	3	3	2	3	2	2	2	2	4	4	4	1	1	1	2	1	3	2
5	1	3	4	3	2	3	2	3	1	4	2	2	4	4	1	1	2	2	1	2	2
6	1	3	4	3	3	3	2	3	3	4	3	2	2	2	1	1	2	2	1	3	2
7	1	3	4	2	3	3	2	2	3	2	3	1	2	2	1	1	2	2	4	3	1
8	3	3	4	1	2	3	2	1	3	4	3	4	2	2	2	2	2	2	3	3	2
9	3	3	4	2	3	3	1	1	1	3	2	4	4	1	2	1	2	2	4	2	2
10	3	3	4	3	2	2	2	2	2	4	3	4	4	2	2	1	2	4	4	2	3
11	3	4	4	3	3	2	2	2	1	1	2	1	4	3	3	1	2	2	1	1	2
12	3	4	4	3	3	3	3	2	1	2	2	1	3	3	3	2	2	2	1	1	2
13	3	4	4	3	2	3	2	3	2	2	3	3	3	3	3	1	2	2	2	2	2
14	3	4	3	2	3	3	2	3	2	3	2	4	3	1	3	1	2	3	1	2	2
15	3	4	4	1	2	3	2	3	3	3	3	3	3	4	3	1	2	3	1	2	3
16	3	4	4	3	3	2	2	2	3	3	3	4	2	4	3	1	2	4	4	2	3
17	3	4	4	2	3	2	3	2	3	3	2	4	1	3	3	1	2	4	4	2	3
18	3	4	3	3	2	2	2	2	2	4	3	4	1	3	2	1	2	3	4	3	3
19	3	4	4	4	2	2	2	2	3	4	3	4	2	4	3	1	3	3	1	3	4
...																					...
100	3	4	4	4	3	3	2	2	1	4	3	4	2	4	2	1	2	3	1	3	3

Table 5
Reduction of conditional attributes.

Serial number	Minimal relative reduction set (ie remaining after reduction)	Strength	Number
1	{C4, C13, C15}	100	3
2	{C9, C12, C19}	100	3
3	{C8, C9, C10, C14}	100	4
4	{C8, C9, C15, C19}	100	4
5	{C9, C10, C17, C20}	100	4
6	{C10, C12, C15, C18}	100	4
7	{C9, C15, C17, C19}	100	4
8	{C4, C14, C17, C18}	100	4
9	{C2, C9, C19, C20}	100	4
10	{C8, C12, C14, C17}	100	4
11	{C10, C12, C15, C19}	100	4
12	{C1, C9, C19, C20}	100	4
13	{C14, C15, C19, C20}	100	4
14	{C1, C9, C10, C14}	100	4
15	{C9, C12, C15, C18}	100	4
16	{C10, C14, C15, C19}	100	4
17	{C2, C9, C10, C18}	100	4
18	{C4, C7, C17, C18}	100	4
19	{C2, C9, C12, C20}	100	4
20	{C1, C14, C17, C19}	100	4
21	{C2, C7, C9, C10}	100	4
22	{C1, C8, C9, C18}	100	4
23	{C4, C7, C12, C17}	100	4
24	{C2, C4, C10, C17, C19}	100	5
25	{C2, C8, C9, C10, C13}	100	5
26	{C1, C4, C9, C14, C18}	100	5
27	{C7, C9, C10, C13, C14}	100	5
28	{C2, C4, C8, C9, C17}	100	5

factors of building energy consumption. The specific division of the discrete points of each factor is shown in Table 2.

The original domain of the 100 power consumption datasets used for attribute reduction of the rough set is discretized into four energy consumption levels, as shown in Table 3. Less than 50 kwh is low energy consumption, which is represented by level 1. From 50 kwh to 80 kwh is medium energy consumption, which is represented by level 2. From 80 kwh to 110 kwh is higher energy consumption, which is represented by level 3. >110 kwh is high energy consumption, which is represented by level 4.

The 20 inputs for building energy consumption $C_1, C_2, C_3, \ldots, C_{20}$ are used as the 20 conditional attributes, and the conditional attribute values are the four coding level values in Table 2. The amount of energy consumption is taken as a decision attribute, and the decision value is the sequence value of the four energy consumption grades determined according to the energy consumption in Table 3. This study considered a data sample as an object and condition variables (the inputs) jointly determined the decision variable (building energy consumption or output). Based on the data collected in Table 1, the decision table for building energy consumption is shown in Table 4.

After the attribute reduction operation, a total of 28 reductions with a strength of 100 are obtained, as shown in Table 5. From the reduction results, C_1, C_2, C_4, C_7, C_8, C_9, C_{10}, C_{12}, C_{13}, C_{14}, C_{15}, C_{17}, C_{18}, C_{19}, and C_{20} repeatedly appear in the multiple minimum relative reduction results, which shows that these 15 conditional attributes (reassigned as Z_1, Z_2, …, Z_{15}) are the core of the decision table. Therefore, through the attribute reduction of the rough set, five miscellaneous and unimportant inputs of building energy consumption, including wind speed, number of floors, building area, building length-to-width ratio and fresh air volume per capita, are eliminated.

4.2. Prediction of building energy consumption

In the short-term prediction of building energy consumption, there are 8176 sets of original data, which are divided into two parts: test data and training data. Among them, 336 sets are used as test data for two test periods in the low-energy season and high-energy season, and the rest are used as training data. In the medium-term prediction of building energy consumption, a total of 341 sets of data are obtained, 30 sets are used for testing, and 311 sets are used for training. For comparison, this study also trained and tested the BP neural network, Elman neural network and fuzzy neural network (FNN). This study used trial and error method to find the optimal structure and parameters of these three neural networks. For example, the optimal structure of BP neural network was a double hidden layer neural network. The number of neuron nodes in the first hidden layer was 10, and the number of neuron nodes in the second hidden layer was 14.

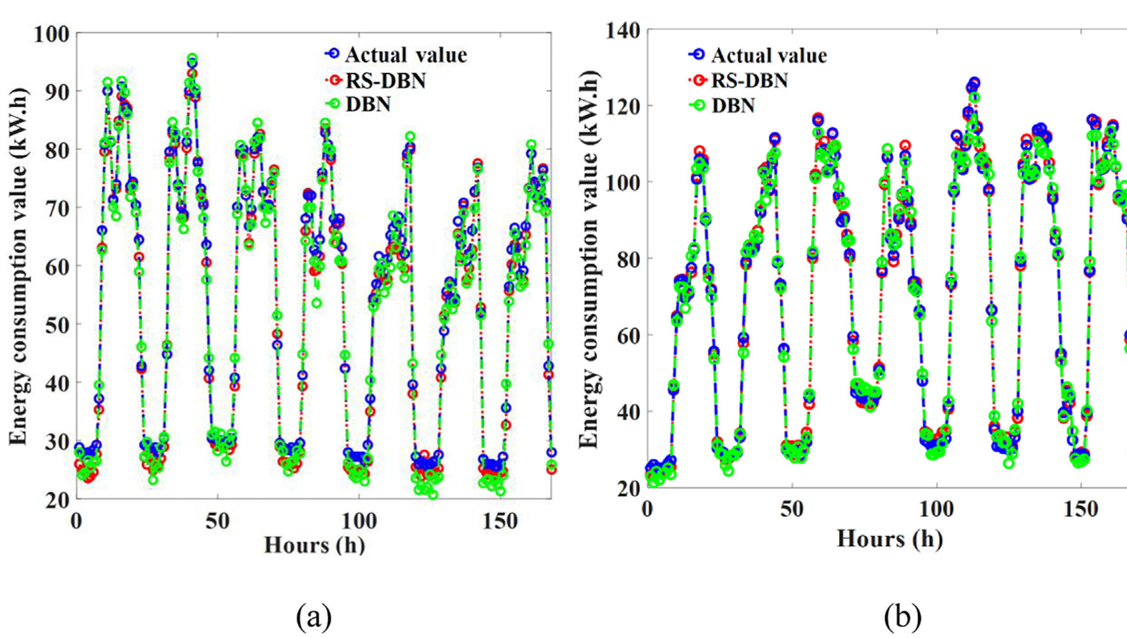

Fig. 9. The prediction of building energy consumption by DBN and RS-DBN for (a) low energy consumption season and (b) high energy consumption season.

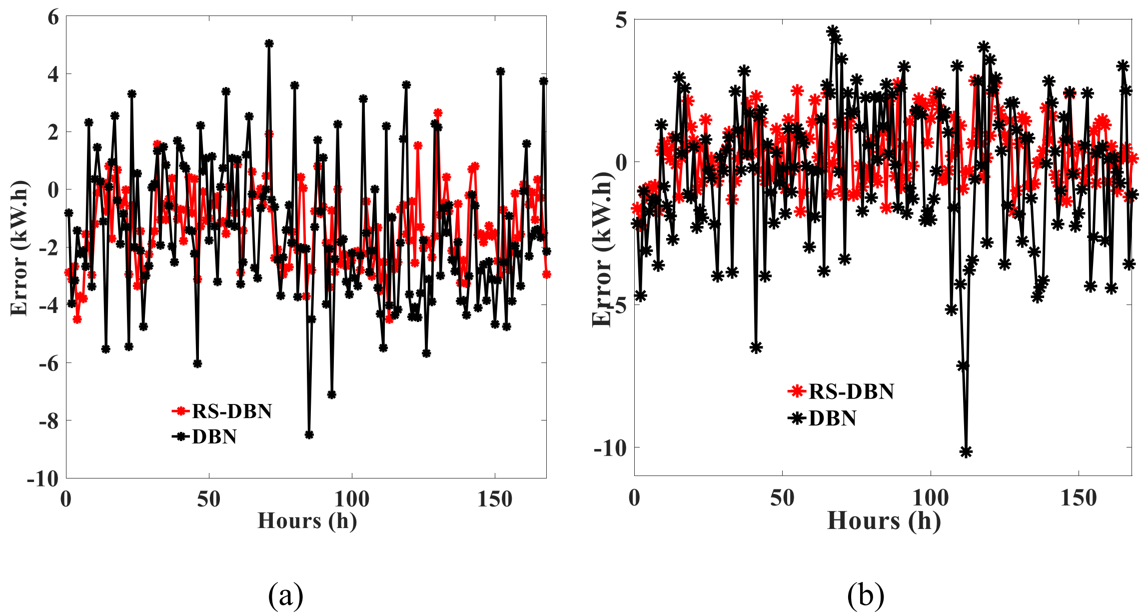

Fig. 10. The prediction error of DBN and RS-DBN for (a) low energy consumption season; (b) high energy consumption season.

Table 6
Quantitative evaluation of DBN model and RS-DBN model.

Prediction model	Low energy consumption season		High energy consumption season	
	MAPE	RMSPE	MAPE	RMSPE
DBN	0.0552	0.0515	0.0484	0.0498
RS-DBN	0.0388	0.0347	0.0291	0.0265

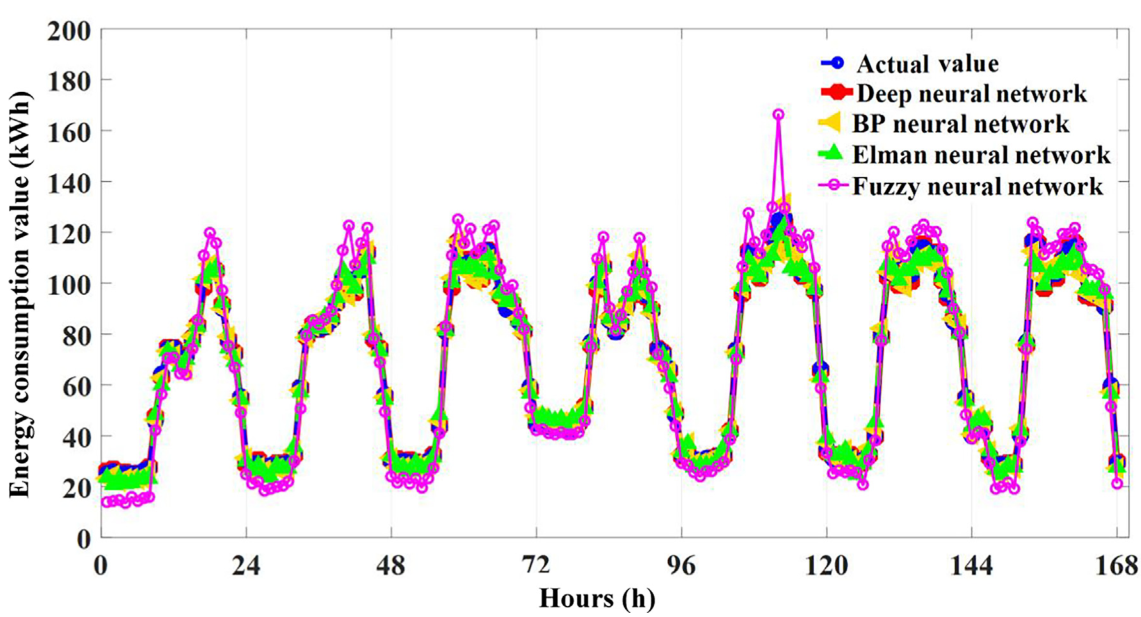

Fig. 11. Comparison of short-term prediction of building energy consumption with the experimental data for the laboratory building in summer.

4.2.1. Short-term prediction of building energy consumption without the rough set

This section investigated the necessity of integrating rough set with neural network. Without using rough sets as the front-end processor of the prediction model, the 20 inputs of building energy consumption were directly used as the inputs of the four neural networks. The test used short-term prediction of building energy consumption for the laboratory building during the low-energy season and the high-energy season. Fig. 9 shows the prediction of the DBN model without rough set and the RS-DBN model. Fig. 10 shows the error of the DBN without rough set and RS-DBN in the two time periods. The results show that the prediction performance of the RS-DBN model is better than that of the DBN model without the rough set.

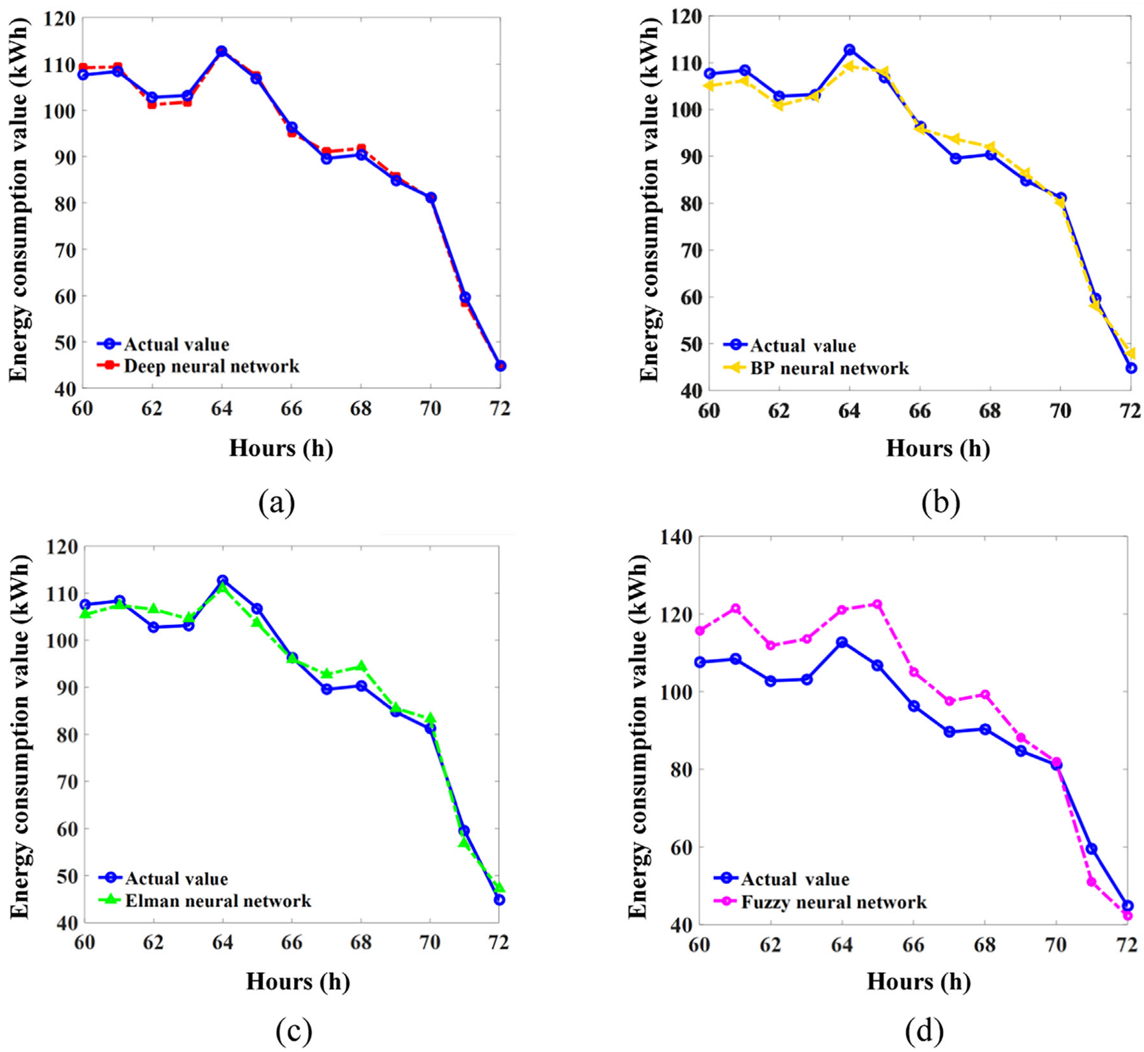

Fig. 12. Comparison of short-term prediction of building energy consumption with the experimental data for the laboratory building in summer from the 60th hour to the 72nd hour.

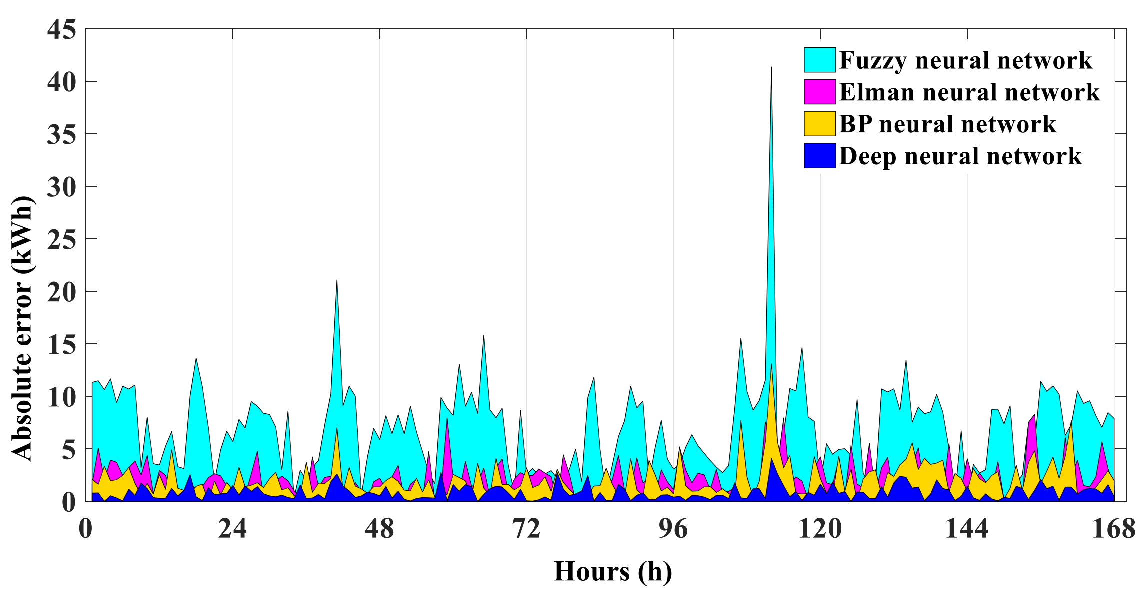

Fig. 13. The absolute error of the four models in short-term prediction of building energy consumption in the summer.

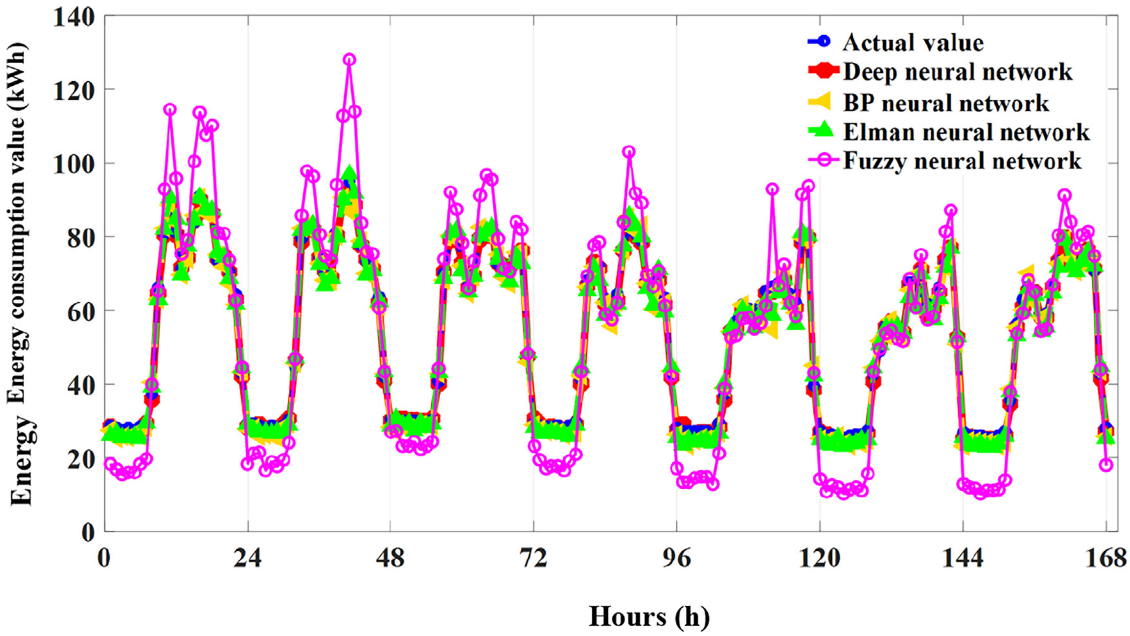

Fig. 14. Prediction results of four models for short-term prediction of building energy consumption in low energy consumption season of the laboratory building.

By calculating average absolute relative error (MAPE) and root mean square relative error (RMSPE),

$$\mathrm{MAPE} = \frac{1}{n}\sum\nolimits_{i=1}^{n}\left|\frac{y_q - y_p}{y_q}\right| \tag{1}$$

$$\mathrm{RMSPE} = \frac{1}{\overline{y_q}}\sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}(y_q - y_p)^2} \tag{2}$$

the prediction accuracy of DBN model and RS-DBN model was quantitatively evaluated. Among them, y_q and y_p represent actual value and predicted value, respectively, and $\overline{y_q}$ represents the average of actual value. The calculation results are shown in Table 6.

For both seasons, the calculated MAPE and RMSPE of the RS-DBN model were smaller than the DBN model. It indicates that without rough set participation, the prediction of building energy consumption by a neural network alone has a lower prediction accuracy than that by the integration of the rough set and a neural network. Therefore, the following tests used the integrated rough set and neural networks.

4.2.2. Short-term prediction of building energy consumption with the rough set

This study conducted the prediction of building energy consumption in the summer, which is the season with high energy consumption in the laboratory building, and spring, which is the season with low energy consumption. After rough sets was used to reduce the influencing factors of building energy consumption, 15 important influencing factors of building energy consumption remained. These remaining factors were used as the inputs of the deep belief neural network, and hourly building energy consumption was output.

From June 23 to June 29 (a total of 168 h, summer), the hourly building energy consumption for the laboratory building is predicted. Fig. 11 compares the predicted hourly energy consumption with the measured one of the laboratory building. The trend of the predicted energy consumption of the four neural networks is generally consistent with that of the measured data, which shows that the combination of the rough set and neural network to predict building energy consumption is effective and feasible. To make clearer and more intuitive comparison between the predictions and experimental data, Fig. 12 presents the results from the 60th hour to the 72nd hour. The comparison shows that the predicted energy consumption of DBN has the highest agreement with the actual energy consumption.

Fig. 13 shows the overall performance of the four models by calculating the absolute error in the short-term prediction of building energy consumption in the summer. The Elman neural network and fuzzy neural network, the deep neural network is more accurate than the BP neural network in predicting the energy consumption of the laboratory building in the summer season.

Spring is representative of the low energy consumption season of the laboratory building. From March 20 to March 26, short-term prediction of building energy consumption was carried out by the four neural networks as shown in Fig. 14. The results show that the prediction of FNN deviated a lot with the experimental data, while the predictions of other three neural networks agreed well with the experimental data. The major difference between the spring and summer in terms of energy consumption is the usage of air-conditioning system. This comparison confirmed that DBN model can accurately accomplish short-term prediction of building energy consumption with/without air-conditioning system.

The experimental results from the 60th hour to the 72nd hour is enlarged in Fig. 14 to further compare the prediction effects of the RS-BP, RS-SRN and RS-DBN in the low energy consumption season. The comparison result is shown in Fig. 15. The predicted energy consumption curve of the RS-DBN has the highest fitting degree with the actual energy consumption curve, and its prediction effect is the best. This result indicates that the RS-DBN prediction model is still more advantageous than other models in short-term prediction of building energy consumption in the low energy consumption season of the laboratory building.

Fig. 16 is a comparison of the overall error performance of the short-term prediction of building energy consumption by the four prediction models during the low energy consumption season of the laboratory building. The area surrounded by the absolute value curve of the error of the RS-DBN model and the X axis is mainly covered by the area representing the errors of the other three models. The overall error performance of the RS-DBN model is the best, and its prediction error for most test samples is at a minimum. In

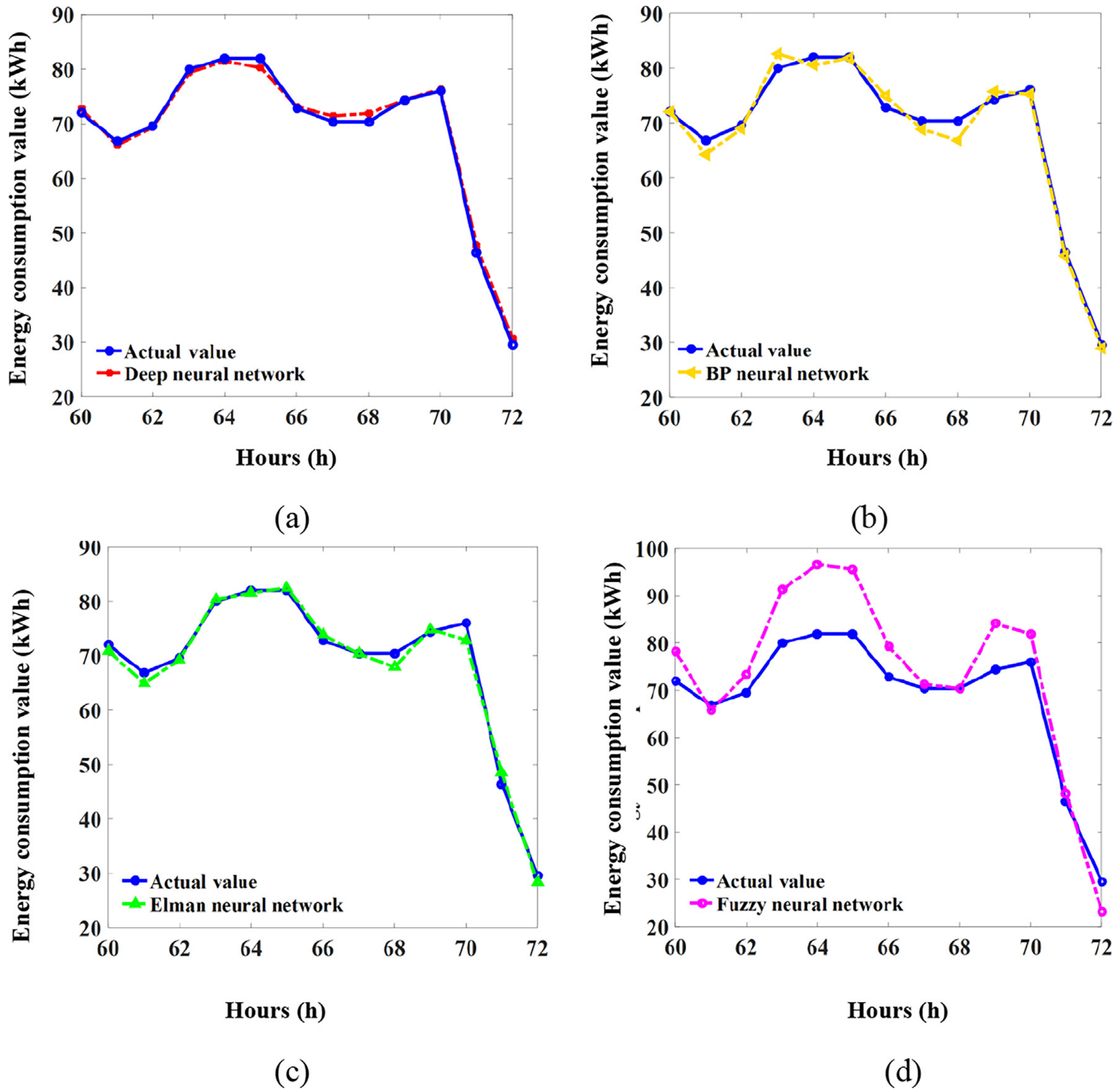

Fig. 15. Partial results of short-term prediction of building energy consumption in low energy consumption season of the laboratory building: (a) RS-DBN; (b) RS-BP; (c) RS-SRN; (d) RS-FNN.

general, the accuracy performance of the four models in the low energy consumption season is roughly similar to that in the high energy consumption season. The RS-DBN model can accurately and effectively predict short-term building energy consumption under the influence of the air conditioning system factors to varying degrees.

4.2.3. Medium-term prediction of building energy consumption with the rough set

The suitability of the RS-DBN prediction model is also investigated for mid-term prediction of building energy consumption. The hourly energy consumption data of the laboratory building are processed into daily energy consumption data, and a 30-day mid-term prediction of building energy consumption is conducted in November (during winter). The prediction results of the four models of the RS-DBN, RS-BP, RS-SRN and RS-FNN for daily building energy consumption in November are shown in Fig. 17. The predictions by the RS-DBN and RS-SRN have a high degree of fitting with the actual energy consumption. Therefore, the RS-DBN and RS-SRN are more suitable than the RS-BP and RS-FNN for mid-term prediction of building energy consumption.

To facilitate a clear comparison of the prediction accuracy of the four models on the medium-term prediction of building energy consumption, a needle chart is used to analyze the absolute error of each model. The absolute errors of the four models for each test data are shown in Fig. 18. Among them, at most test sample points, the length of the needles representing the RS-BP or RS-FNN is the longest and the fluctuation range is large. The needle length of the RS-DBN model is basically the shortest and its range of fluctuation is small. The RS-DBN model has the smallest absolute error in the mid-term prediction of building energy consumption and its error has the least variations. Table 7 further shows the relative error of the RS-DBN model for the mid-term prediction of building energy consumption. The maximum relative error of the RS-DBN model does not exceed 7%. Therefore, the RS-DBN model is not only suitable for short-term prediction of building energy consumption, but also accurate and effective for mid-term prediction of building energy consumption.

5. Discussions

This study used considerably high percentage of data for training because it is dependent on the nature of the deep neural net-

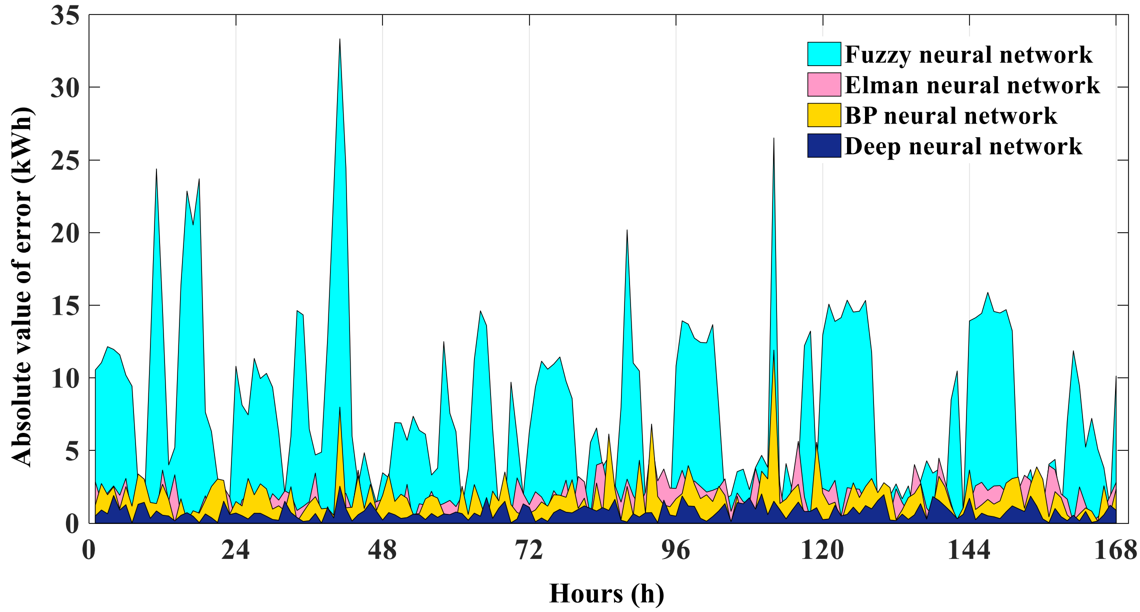

Fig. 16. The overall error performance of the four models in the prediction experiment of low energy consumption season.

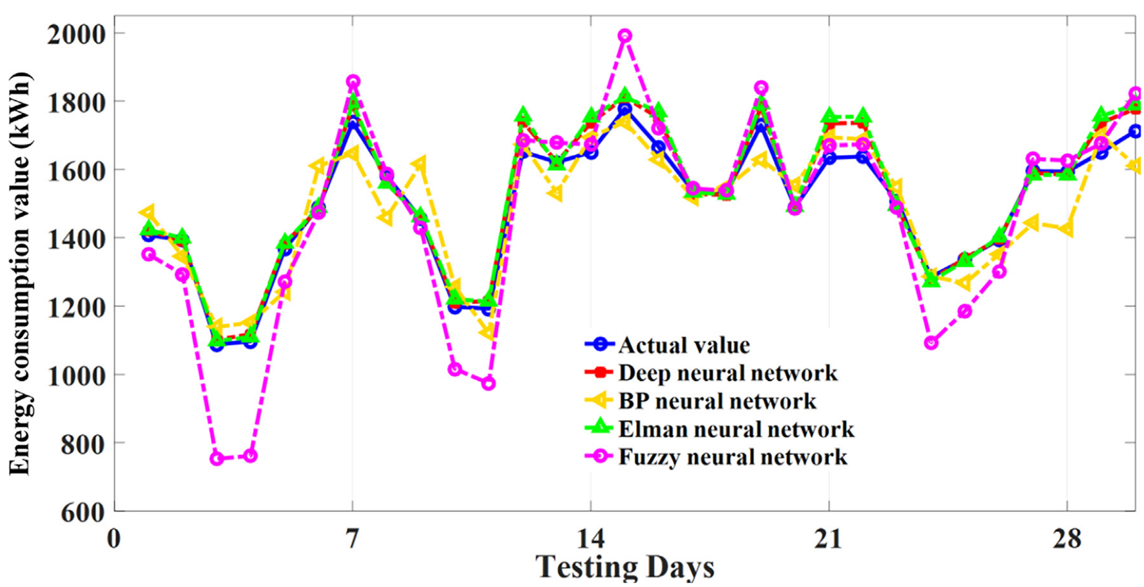

Fig. 17. The prediction results of the four models in the medium-term prediction experiment of building energy consumption.

work. The main training method of deep belief neural network is unsupervised learning, which requires a large amount of training data for the algorithm itself to extract the characteristics of the data. In contrast, the shallow neural network generally uses supervised learning for training method, which requires less amount of training data. For example, the two studies [25] and [21] used considerably high percentage of data for training. Li et al. [21] used >90% of the data for training.

One can notice that this study deducted the relevant input variables by using the dataset of public buildings and applied the results for a laboratory building in training and testing the deep neural network. One reason is that the civil buildings refer to non-residential buildings where people carry out various public activities, such as office buildings, commercial buildings, tourist buildings, science, education, culture and health buildings, and transportation buildings, etc. The laboratory building belongs to civil public buildings. In terms of schedule, the open time of those buildings are generally from 7:00 am to 10:00 pm. In terms of building energy consumption, the laboratory building is similar with other civil public buildings as most of its energy is consumed by the HVAC system, computers, experimental equipment, and lighting. Another reason is that the required data for rough set theory is its diversity but not quantity. Therefore, this study collected the data of 100 civil public buildings for only one hour each. However, the data used for training and testing the deep neural network requires long-term monitoring, the data collection is difficult in those civil public buildings. For a laboratory building located in a university, the long-term monitoring of data in a controllable range could be realized.

The “feature” in this study refers to the implicit connections between the input and the output data of the deep neural network. These connections will directly affect the output information after

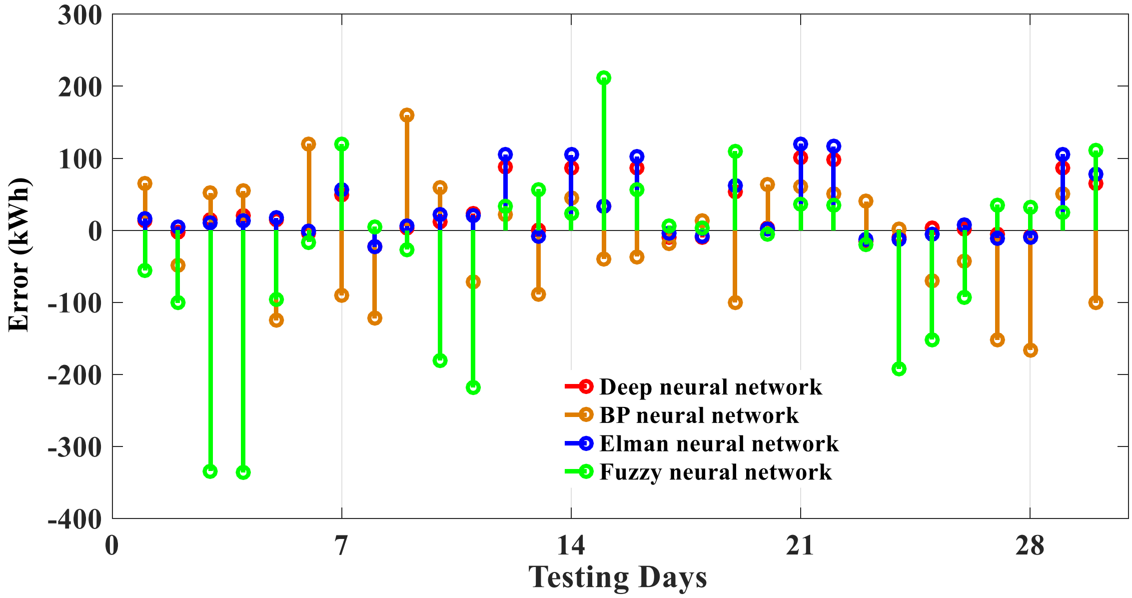

Fig. 18. The absolute errors of the four models in the medium-term prediction experiment of building energy consumption.

Table 7
Relative error analysis of RS-DBN model in the medium-term prediction experiment of building energy consumption.

Days	Actual energy consumption value (kwh)	RS-DBN predicted energy consumption value (kwh)	Relative error of RS-DBN model in the medium-term prediction (%)
Day 1	1407.2	1419.8	0.90%
Day 2	1394	1391.4	0.19%
Day 3	1087.6	1102.1	1.33%
Day 4	1097.2	1117.9	1.89%
Day 5	1367.6	1382.2	1.07%
Day 6	1489.2	1485.6	0.24%
Day 7	1738	1787.8	2.87%
Day 8	1581.6	1559.2	1.42%
Day 9	1455.6	1459.2	0.25%
Day 10	1197.6	1208.9	0.94%
Day 11	1192.4	1215.3	1.92%
Day 12	1652	1740	5.33%
Day 13	1621.2	1621.1	0.01%
Day 14	1650.4	1737.4	5.27%
Day 15	1778.4	1811.1	1.84%
Day 16	1665.2	1751.1	5.16%
Day 17	1537.6	1528.6	0.59%
Day 18	1535.2	1524.9	0.67%
Day 19	1730	1782.9	3.06%
Day 20	1489.6	1493	0.23%
Day 21	1633.6	1734.2	6.16%
Day 22	1638.4	1736.8	6.01%
Day 23	1506.8	1491.9	0.99%
Day 24	1285.2	1275.8	0.73%
Day 25	1337.6	1340.4	0.21%
Day 26	1394.8	1396.6	0.13%
Day 27	1595.2	1589.5	0.36%
Day 28	1594	1585.4	0.54%
Day 29	1650.4	1736.7	5.23%
Day 30	1712.4	1777.9	3.83%

passing through layer-by-layer and bottom-up unsupervised learning. This study did not discuss the actual feature extraction because deep learning used an algorithm for learning features automatically. The training method of the deep neural network was unsupervised. It used uncalibrated data to train the parameters of each layer spontaneously and does not require artificial intervention in the process of extracting features.

6. Conclusions

This paper integrated the rough sets and deep learning to establish a prediction model for building energy consumption. The model uses rough set theory as the front-end processor for deep learning. The attribute reduction method based on the genetic algorithm in rough set theory is used to reduce the miscellaneous influencing factors of building energy consumption. The model uses a deep neural network as the post information recognition system. The remaining influencing factors of building energy consumption after rough set reduction were used as the input of the DBN, and the building energy consumption are output from the DBN. This investigation tested the developed model by comparing its performance with BP, Elman and fuzzy neural networks in both short-term and medium-term predictions of building energy consumption. The results led to the following conclusions:

- The implementation of rough set theory was able to eliminate redundant influencing factors of building energy consumption. The DBN with reduced number of inputs had improved accuracy in building energy simulation.
- The DBN had more accurate prediction of either short-term or long-term building energy consumption than the shallow neural networks such as BP, Elman and fuzzy neural networks.
- This study presented a procedure to make use of machine learning for building energy simulation. The developed prediction model has great potential to predict the energy consumption of various types of buildings on different time scales. It can bring more accurate prediction results, which can enable energy demand managers to better schedule and control the fluctuating power supply for demand management continuously in real time.

CRediT authorship contribution statement

Lei Lei: Funding acquisition, Project administration, Supervision, Conceptualization, Writing - review & editing, Resources, Methodology. **Wei Chen:** Writing - original draft, Formal analysis, Investigation, Data curation, Software, Validation. **Bing Wu:** Writing - review & editing. **Chao Chen:** Data curation. **Wei Liu:** Funding acquisition, Supervision, Conceptualization, Writing - review & editing, Formal analysis.

Declaration of Competing Interest

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

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant No.: 51708146), Guangxi Natural Science Foundation (Grant No.: 2018GXNSFAA281283), Guangxi Science and Technology Project (Grant No.: Guike AD18281046). Wei Liu acknowledges the financial support from Energimyndigheten (Swedish Energy Agency, grant No. 50057-1).

References

[1] H. Zhong, J. Wang, H. Jia, Y. Mu, S. Lv, Vector field-based support vector regression for building energy consumption prediction, Appl. Energy. 242 (2019) 403–414.
[2] C.D. Li, Z.X. Ding, D.B. Zhao, et al., Building energy consumption prediction: an extreme deep learning approach, Energies 10 (10) (2017) 1–20.
[3] K. Amasyali, N.M. El-Gohary, A review of data-driven building energy consumption prediction studies, Renew. Sustain. Energy Rev. 81 (2018) 1192–1205.
[4] C. Fan, F.u. Xiao, Y. Zhao, A short-term building cooling load prediction method using deep learning algorithms, Appl. Energy 195 (2017) 222–233.
[5] K. Li, C. Hu, G. Liu, W. Xue, Building's electricity consumption prediction using optimized artificial neural networks and principal component analysis, Energy Build. 108 (2015) 106–113.
[6] G. Qiang, T. Zhe, D. Yan, Z. Neng, An improved office building cooling load prediction model based on multivariable linear regression, Energy Build. 107 (2015) 445–455.
[7] H. Pombeiro, R. Santos, P. Carreira, et al., Comparative assessment of low-complexity models to predict electricity consumption in an institutional building: Linear regression vs. fuzzy modeling vs. neural networks, Energy Build. 146 (2017) 141–151.
[8] M.A. Jallal, A.G. Vidal, A.F. Skarmeta, A hybrid neuro-fuzzy inference system-based algorithm for time series forecasting applied to energy consumption prediction, Appl. Energy 268 (2020) 114977, Energy Build. 111 (2016) 380-392.
[9] Z.W. Li, Y.M. Han, P. Xu, Methods for benchmarking building energy consumption againstits past or intended performance: an overview, Appl. Energy 124 (7) (2014) 325–334.
[10] J. Braun, N. Chaturvedi, An inverse gray-box model for transient building load prediction, HVAC&R Res. 8 (1) (2002) 73–99.
[11] B. Dong, C. Cao, S.E. Lee, Applying support vector machines to predict building energy consumption in tropical region, Energy Build. 37 (5) (2005) 545–553.
[12] L. Robert, R. Uwe, Comparison between support vector regression and grey system theory based models for adaptive load forecasting of building thermal energy consumption, EG-ICE 2014, European Group for Intelligent Computing in Engineering-21st International Workshop: Intelligent Computing in Engineering 2014. 2014.
[13] J.J. Guo, J.Y. Wu, R.Z. Wang, A new approach to energy consumption prediction of domestic heat pump water heater based on grey system theory, Energy Build. 43 (6) (2011) 1273–1279.
[14] Y. Wei, L. Xia, S. Pan, J. Wu, X. Zhang, M. Han, W. Zhang, J. Xie, Q. Li, Prediction of occupancy level and energy consumption in office building using blind system identification and neural networks, Appl. Energy. 240 (2019) 276–294.
[15] B.B. Ekici, U.T. Aksoy, Prediction of building energy consumption by using artificial neural networks, Adv. Eng. Softw. 40 (5) (2009) 356–362.
[16] A. Rahman, V. Srikumar, A.D. Smith, Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks, Appl. Energy 212 (2018) 372–385.
[17] K. Li, H. Su, J. Chu, Forecasting building energy consumption using neural networks and hybrid neuro-fuzzy system: a comparative study, Energy Build. 43 (10) (2011) 2893–2899.
[18] P. Domingos, A few useful things to know about machine learning, Commun. ACM 55 (10) (2012) 78–87.
[19] A. Liaw, M. Wiener, Classification and regression by randomForest, R News. 2 (3) (2002) 18–22.
[20] S. Naji, A. Keivani, S. Shamshirband, U.J. Alengaram, M.Z. Jumaat, Z. Mansor, M. Lee, Estimating building energy consumption using extreme learning machine method, Energy. 97 (2016) 506–516.
[21] C.D. Li, Z.X. Ding, D.B. Zhao, J. Yi, G. Zhang, Building energy consumption prediction: an extreme deep learning approach, Energies. 10 (10) (2017) 1525.
[22] K. Shankar, A.R.W. Sait, D. Gupta, S.K. Lakshmanaprabu, A. Khanna, H.M. Pandey, Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model, Pattern Recogn. Lett. 133 (2020) 210–216.
[23] I.R.I. Haque, J. Neubert, Deep learning approaches to biomedical image segmentation, Inf. Med. Unlocked 18 (2020) 100297.
[24] A. Kim, Y. Yang, S. Lessmann, T. Ma, M.-C. Sung, J.E.V. Johnson, Can deep learning predict risky retail investors? A case study in financial risk behavior forecasting, Eur. J. Operation. Res. 283 (1) (2020) 217–234.
[25] M.A. Amanullah, R.A.A. Habeeb, F.H. Nasaruddin, A. Gani, E. Ahmed, A.S.M. Nainar, N.M. Akim, M. Imran, Deep learning and big data technologies for IoT security, Comput. Commun. 151 (2020) 495–517.
[26] H. Fujiyoshi, T. Hirakawa, T. Yamashita, Deep learning-based image recognition for autonomous driving, IATSS Res. 43 (4) (2019) 244–252.
[27] P. Hähnel, J. Mareček, J. Monteil, F. O'Donncha, Using deep learning to extend the range of air pollution monitoring and forecasting, J. Comput. Phys. 408 (2020) 109278.
[28] F. Zhang, C. Deb, S.E. Lee, J. Yang, K.W. Shah, Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique, Energy Build. 126 (2016) 94–103.
[29] F. Rodrigues, C. Cardeira, J.M.F. Calado, The daily and hourly energy consumption and load forecasting using artificial neural network method: a case study using a set of 93 households in Portugal, Energy Proc. 62 (2014) 220–229.
[30] P. Lusis, K.R. Khalilpour, L. Andrew, et al., Short-term residential load forecasting: Impact of calendar effects and forecast granularity, Appl. Energy 205 (2017) 654–669.
[31] M. Ghiassi, D.K. Zimbra, H. Saidane, Medium term system load forecasting with a dynamic artificial neural network model, Electr. Power Syst. Res. 76 (5) (2006) 302–316.
[32] M. Cunkas, A.A. Altun, et al., Long term electricity demand forecasting in Turkey using artificial neural networks, Energy Source Part B: Econ. Plann. Policy. 5 (2010) 279–289.
[33] Z. Pawlak, Rough set, Int. J. Comput. Inform. Sci. 11 (2) (1982) 341–356.
[34] Q.L. Zhou, H.M. Mo, Y. Deng, A new divergence measure of pythagorean fuzzy sets based on belief function and its application in medical diagnosis, Mathematics 8 (1) (2016) 389–411.

[35] A. Skowron, H.S. Nguyen, Boolean reasoning scheme with some applications in data mining, in: European Conference on Principles of Data Mining & Knowledge Discovery, Springer-Verlag, 1999, pp. 107–115.
[36] X. Zhu, Y. Qian, B. Sun, et al., Image Inpainting Forensics Algorithm Based on Deep Neural Network, Acta Optica Sinica. 38 (11) (2018) 105–113.
[37] T. L∅ken, Vibration Analysis using Rosetta-A Practical Application of Rough Sets, The Norwegian University of Science and Technology, 1998.
[38] J. Sunarso, S.S. Hashim, J.Y.J. Yeo, J.J. Chew, MATLAB-based project assessment in process modelling unit: A case study from Swinburne University of Technology Sarawak Campus, Educ. Chem. Eng. 33 (2020) 17–26.
[39] J.P. Liu, Building Physics, China Architecture & Building Press, Beijing, 2009.