Energy & Buildings 262 (2022) 111718

Contents lists available at ScienceDirect

Energy & Buildings

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

Building energy prediction using artificial neural networks: A literature survey

Chujie Lu [a], Sihui Li [b], Zhengjun Lu [c,*]

[a] School of Computers, Guangdong University of Technology, Guangzhou 510006, PR China
[b] College of Civil Engineering, Hunan University, Changsha 410082, PR China
[c] Defense Engineering Institute, AMS, PLA, Beijing 100850, PR China

ARTICLE INFO

Article history:
Received 15 April 2021
Revised 29 October 2021
Accepted 20 November 2021
Available online 26 November 2021

Keywords:
Building energy prediction
Artificial neural networks
Deep learning
Smart buildings

ABSTRACT

Building Energy prediction has emerged as an active research area due to its potential in improving energy efficiency in building energy management systems. Essentially, building energy prediction belongs to the time series forecasting or regression problem, and data-driven methods have drawn more attention recently due to their powerful ability to model complex relationships without expert knowledge. Among those methods, artificial neural networks (ANNs) have proven to be one of the most suitable and potential approaches with the rapid development of deep learning. This survey focuses on the studies using ANNs for building energy prediction and provides a bibliometric analysis by selecting 324 related publications in the recent five years. This survey is the first review article to summarize the details and applications of twelve ANN architectures in building energy prediction. Moreover, we discuss three open issues and main challenges in building energy prediction using ANNs regarding choosing ANN architecture, improving prediction performance, and dealing with the lack of building energy data. This survey aims at giving researchers a comprehensive introduction to ANNs for building energy prediction and investigating the future research directions when they attempt to implement ANNs to predict building energy demand or consumption.

1. Introduction

According to the International Energy Agency, the buildings and buildings construction sectors combined are responsible for over one-third of global final energy consumption and nearly 40% of total direct and indirect CO2 emissions [1]. Due to the urbanization process in the developing countries and the climate change, such as global warming and extreme weather events, building energy consumption will keep growing. To solve this problem, many energy conservation measures during design have been proposed to improve the energy efficiency of buildings, i.e. better building envelope design [2], Meanwhile, with the rapid development of the Internet of Things (IoT), the deployment of large-scale sensing, big data analytics, and advanced control has been spurring the evolution of buildings from simple to smart [3]. Much real-time building operation data is collected and processed in smart buildings, go beyond measuring temperature and humidity to sensing air quality, light intensity, and occupancy information [4]. IoT technologies would be another solution through implementing advanced strategies to achieve energy-efficient buildings, including predictive and occupant-centric control optimization for lighting systems and heating, ventilation, and air-conditioning (HVAC) systems [5], thermal energy storage strategy [6], renewable energy integration [7], and smart grid management [8].

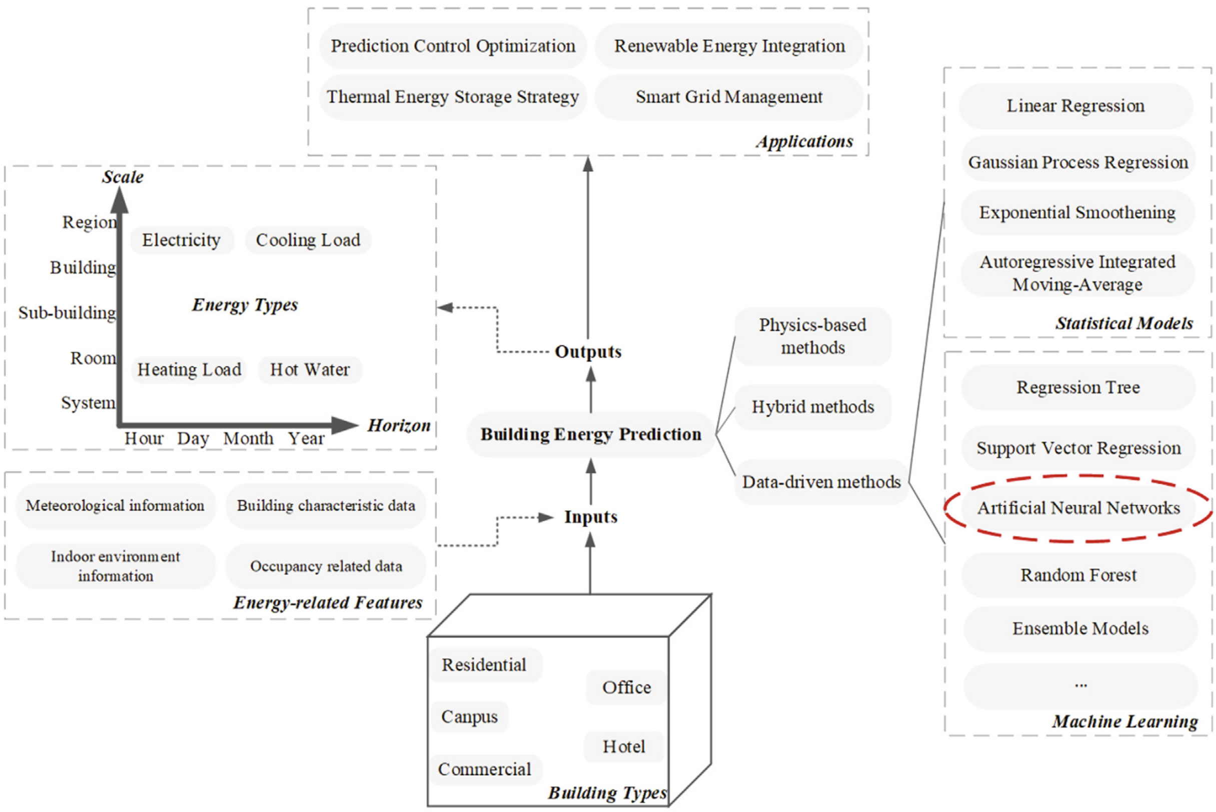
Abbreviations: ABC, Artificial bee colony; ANN, Artificial neural networks; ARIMA, Autoregressive integrated moving average; BA, Bat algorithm; CS, Cuckoo Search; DA, Dragonfly algorithm; DE, Differential evolution algorithm; EA, Evolutionary algorithm; ELM, Extreme learning machine; ENN, Elman neural networks; ESN, Echo state networks; FFNN, Feed forward neural networks; FFOA, Fruit fly optimization algorithm; GA, Genetic algorithm; GAN, Generative adversarial network; GPR, Gaussian process regression; GRU, Gated recurrent units; HVAC, Heating, ventilation, and air-conditioning; ICA, Imperialist competitive algorithm; IoT, Internet of Things; LSTM, Long short-term memory; MLP, Multilayer perceptron; MLR, Multiple linear regression; NARX, Nonlinear autoregressive neural network with exogenous inputs; PSO, Particle swarm optimization; RBF, Radial basis function; RBM, Restricted Boltzmann machine; RF, Random forest; RNN, Recurrent neural networks; SCOA, Sine cosine optimization algorithm; SOS, Symbiotic organism search; SVR, Support vector regression; TSBO, Teaching–learning-based optimization; WNN, Wavelet neural networks.

* Corresponding author.
E-mail address: LUZJ@sohu.com (Z. Lu).

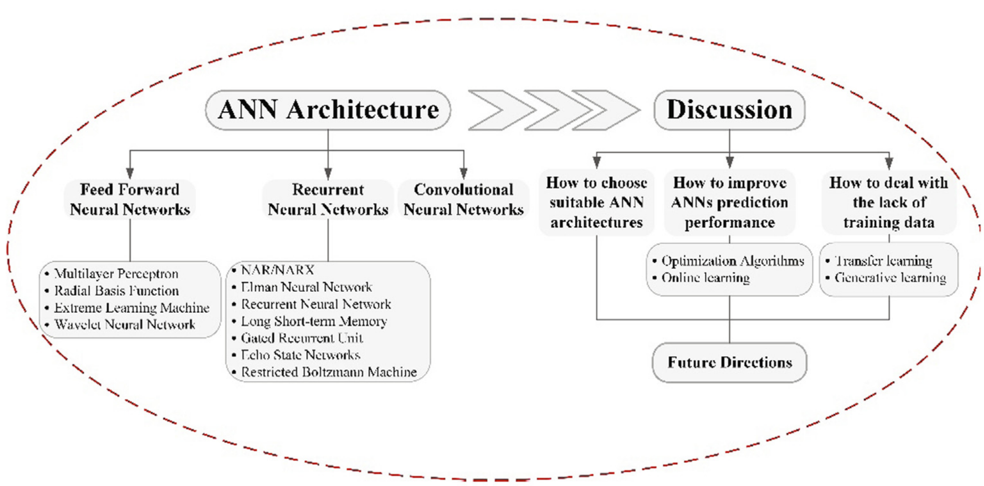
https://doi.org/10.1016/j.enbuild.2021.111718
0378-7788/

Building energy prediction is not only an important evaluation tool of energy-saving potential during building design and retrofit but also an essential component of smart buildings, illustrated in Fig. 1(a). The definition of building energy could refer to [9,10], which has the characteristics of complexity, dynamics, and nonlinearity. Building energy could be divided roughly into cooling load, heating load, hot water, lighting load, plug load, and overall electricity in terms of energy end-use, while they also could be divided as residential, office, campus, industry, and commercial in terms of building type. Common approaches to predict building energy performance include physics-based method, hybrid method, and data-driven method. Among those approaches, the data-driven method has attracted more and more attention in recent years because it is less time-consuming, no requirements of expertise knowledge, easy-deployed in practice and could obtain more accurate prediction performance.

The data-driven method could be considered as a black box, ignoring the internal detailed relationships of the heat and mass transferring in buildings. As large amounts of smart meters are deployed and increasing amounts of data are collected, it becomes possible to learn the pattern of building energy through some advanced and complicated data-driven models in practice. Build-



(a)



(b)

Fig. 1. Illustration showing (a) the focus of this review article within the domain of building energy systems, (b) The framework of this review article.

ing energy prediction is essentially a time series forecasting problem, which is the process of predicting or judging future phenomena by analyzing some series of historical observation data arranged in chronological order [11,12]. The data-driven method to predict building energy could be univariate or multivariate based on the number of series used in forecasting [12]. In the univariate prediction, an appropriate model depends on carefully analyzing historical energy consumption, and then the model is used to predict the future values of energy consumption. While in the multivariate prediction, in addition to the energy consumption series, the model is developed by values of one or more additional time series, called the energy-related features. Those features should be related to energy consumption and could help improve the prediction performance [10], including meteorological information (such as ambient temperature, humidity, wind speed, solar radiation, etc.), indoor environmental information (such as indoor temperature, indoor humidity, etc.), occupancy information (such as the number of occupants and type of occupant behaviors, etc.) and time index (such as hour of the day, weekday or weekend, etc.). There are some other related features used to enhance the prediction performance, including building characteristic data (such as surface area, roof area, wall area, glazing area, window-wall ratio, heat transfer coefficient of building envelopes, etc.) and socioeconomic information (such as income, electricity price, GDP, etc.).

The early data-driven method was mainly statistical models, such as Multiple linear regression (MLR) [13], Gaussian process regression (GPR). Exponential smoothening, and Autoregressive integrated moving average (ARIMA) [14]. The statistical models do well for modeling stationary and linear series, while they may perform poorly when dealing with time series with stochastic behavior like building energy data. Nowadays, machine learning models are a more popular option for building energy prediction, such as Regression tree [15], Random forest (RF) [15,16], Support vector regression (SVR) [17,18,19], Artificial neural networks (ANNs), and some ensemble models [20,21]. Among those machine learning models in the existing studies, ANN is used more frequently in building energy prediction [22]. Furthermore, with the rapid development of deep learning, ANNs are considered as the model with great potential for accurate energy prediction [23,24]. ANNs are a type of intelligent information processing method inspired by the biological neural network, which is advantageous in the strong ability to represent and model the nonlinear relationships between inputs and outputs.

In particular, many ANN architectures have been applied in building energy prediction, which is generally divided into three main categories, shown in Fig. 1(b). In Feedforward neural networks (FFNNs), signals are transmitted through connections from the input layer to the output layer, passing through neurons in one or more hidden layers with no memory or feedback connections [25]. Multilayer perceptron (MLP), Radial basis function (RBF) networks, Extreme learning machine (ELM), and Wavelet neural network (WNN) belong to main types of popular FFNN. In Recurrent neural networks (RNNs), connections between layers form a cycle that creates an internal state in the network which exhibits dynamic temporal behavior [26], including nonlinear autoregressive neural network with exogenous inputs (NARX), Elman neural network (ENN), Long short-term memory (LSTM), Gated recurrent unit (GRU), Echo state network (ESN), and Restricted Boltzmann machine (RBM). In Convolutional neural networks (CNNs), there are two fundamental operations, convolution and pooling, to extract features automatically. The details of these three architectures (twelve types) and in building energy prediction will be reviewed and some improvement approaches of ANNs prediction performance also will be summarized in this survey.

1.1. Relevant review articles

Data-driven building energy prediction has drawn more attention and there have been many relevant review articles published in the last few years. Note that these review articles are either different from the scope of this article or not as comprehensive and up-to-date as this survey. Sun et al. [10] summarized the general procedure for data-driven building energy prediction and introduced the updating strategies for multi-step building energy prediction. The properties, the uses, and the limitations of existing building energy datasets and data collection platforms were compared in [27]. Various data-driven models and ensemble methods for building energy prediction were reviewed in [9,28,29,30,31,32]. Amasyali et al. [30] further discussed different models in terms of different temporal granularities, building types, and energy consumption types. Wei et al. [32] demonstrated the practical applications of data-driven models in building energy analysis and provided instructive suggestions for different users throughout the whole building life cycle. The key application areas of meter data in the smart grid (including buildings) were discussed in [33]. Especially, Roman et al. [34] gave a general insight of ANNs applied in building energy prediction, including different cases, sample generation, architectures, and training and testing processes. Mohandes et al. [35] reviewed several applications utilizing ANNs in building energy analysis. Table 1 concludes the ANN types mentioned in those review articles and we can find two research gaps. First, most review articles did not introduce ANN types comprehensively, failing to present the important trend of choosing ANN architectures or types in building energy prediction in the recent years. Second, no review article introduces the building energy applications of different ANN architectures in detail and can give effective guidance to choose suitable ANN architectures or types.

1.2. Objectives and review structure

In this article, we aim at conducting a comprehensive literature survey of building energy prediction using ANN, the method most favored by researchers in recent years. The focus of this survey within the domain of building energy systems is illustrated in Fig. 1(a). To be specific, the objectives include (1) providing a systematic literature survey in the past five years and an analysis to show the research trends of ANN applications in building energy prediction; (2) introducing various ANN architectures and types applied in building energy prediction; (3) analyzing the obstacles in current research and investigate the future research directions for researchers.

The main contributions of this survey can be summarized as follows: (1) Show the rising attention on ANNs, especially RNNs, in building energy prediction through a comprehensive bibliometric analysis in the past five years. The research trend and the journal distribution also are presented in detail. (2) Introduce three main architectures (twelve types) of ANNs and illustrate how the related studies apply them to predict building energy performance. (3) Discuss the strategies to solve three following obstacles when predicting building energy performance using ANNs: i. How to choose suitable ANN architecture for building energy prediction. ii. How to improve the performance of ANNs for building energy prediction. iii. How to deal with the lack of energy training data.

As shown in Fig. 1(b), the paper will be organized as follows: Section 2 conducts a comprehensive bibliometric analysis about building energy prediction using ANN. Section 3 provides introductions of various ANN architectures and types, and their applications in different building cases. Section 4 provides some suggestions and recommendations regarding choosing ANN architecture. Section 5 concludes some improvement measures pro-

Table 1
Summary of ANN architectures mentioned in the relevant review articles.

Reference	Year	ANN architectures											
		MLP	RBF	ELM	WNN	NARX	ENN	RNN	LSTM	GRU	ESN	RBM	CNN
[28]	2017	√	√		√		√						
[29]	2017	√	√										
[32]	2018	√						√					
[30]	2018	√	√										
[30]	2018	√	√			√							
[31]	2019	√	√	√		√		√					
[35]	2019	√	√			√		√					
[36]	2019	√			√			√				√	
[34]	2020	√	√			√		√					
[22]	2020	√											
[10]	2020	√						√	√				

posed for ANN applications in building energy prediction in the existing studies. Section 6 discusses two promising techniques for solving the lack of energy data. The conclusion of this literature survey is given in Section 7.

2. Bibliometric analysis in building energy prediction using ANNs

In this section, we will provide a comprehensive overview of the existing studies in building energy prediction using ANNs. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology was applied for the survey, which is a framework to guide systematic reviews and *meta*-analyses [37]. The related publications were found in Web of Science, a well-established and acknowledged database, using the search strings (such as building energy prediction/forecasting, building load prediction/forecasting, building electricity prediction/forecasting, and ANN). We chose these publications in English published between 2016 and 2020. The PRISMA flowchart of this review is shown in Fig. 2. A total of 2154 publications were obtained at first, and 324 publications were selected and included at last through screening and eligibility examination.

Fig. 3 shows the number of publications and the times cited from 2016 to 2020. The number of publications was at a relatively low level before 2017, while it has been increasing rapidly since 2018 and there were up to 100 studies published in 2020. This result is reasonable because deep learning has been rapidly devel-

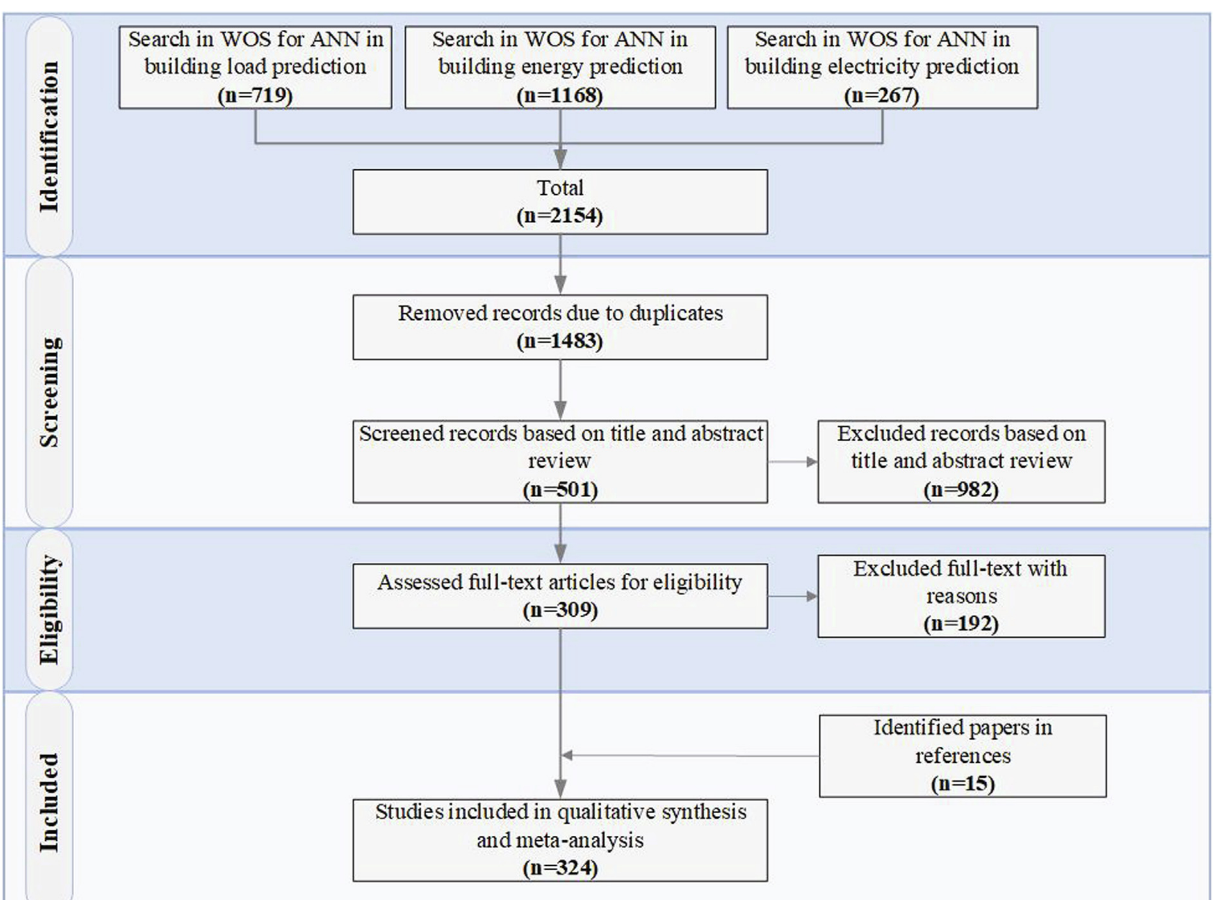


Fig. 2. The PRISMA flowchart of this survey.

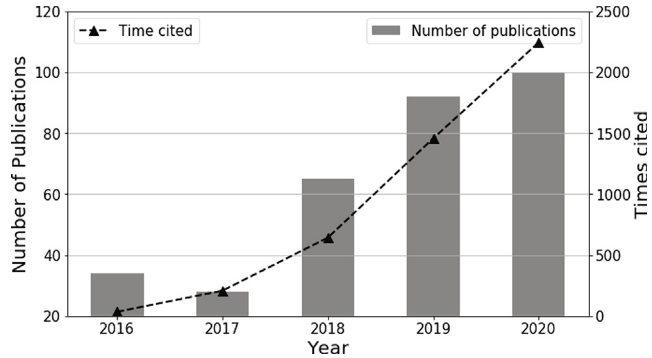


Fig. 3. The number of publications and their times cited in building energy prediction using ANN per year.

oped and applied in many research fields. In summary, researchers have paid more and more attention on building energy prediction using ANNs.

The selected publications include 275 journal articles and the follow-up literature survey mainly focuses on these journal articles. It is important for the interested researchers to know where they can publish their findings. Table 2 depicts the journals of the selected publications ranked by percentages. The journals in the fields of energy engineering and building engineering account for the majority and *Energy and Buildings*, *Energy*, *Energies*, and *Applied Energy* are the top 4 most popular journals. The journals in the fields of automation and computer engineering also appear, such as *IEEE Access*, *Engineering with Computers*, *Applied Soft Computing*, and *Automation in Construction*. The journals in the field of power systems also appear, such as *IEEE Transactions on Smart Grid*. The phenomenon occurs because building energy prediction is a cross-discipline and an important component of smart buildings and smart energy systems, attracting researchers from various fields.

3. ANN architectures in building energy prediction

In this section, we start by describing the general definition and formulation of building energy prediction for ease to understand, and a brief introduction on training and testing of ANNs is presented. Then, we present the taxonomy of the different ANNs and describe in detail the thirteen different ANN architectures and their building energy applications in the selected journal articles.

Table 2
Distribution of the journal articles on building energy prediction using ANN.

Ranking	Journal Title	Number	Percentage(%)
1	*Energy and Buildings*	43	15.64
2	*Energy*	34	12.36
3	*Energies*	29	10.55
4	*Applied Energy*	26	9.45
6	*IEEE Access*	12	4.36
6	*Sustainable Cites and Society*	11	4.00
8	*Applied Science Basel*	7	2.55
8	*Building Simulation*	7	2.55
10	*Applied Thermal Engineering*	5	1.82
11	*Engineering with Computers*	4	1.45
11	*IEEE Transactions on Smart Grid*	4	1.45
11	*Sustainability*	4	1.45
14	*Applied Soft Computing*	3	1.09
14	*Journal of Building Engineering*	3	1.09
14	*Renewable Energy*	3	1.09
14	*Renewable sustainable energy reviews*	3	1.09
14	*Automation in Construction*	3	1.09

3.1. Theoretical background

First, the general definition and formulation of building energy prediction will be displayed. For the univariate prediction, a historical series of building energy, $X^0 = [x_1, x_2, \cdots, x_{T-1}]$, is given, and for the multivariate prediction, a $M-$dimensional series, $X = [X^0, X^1, \cdots, X^{M-1}]$, is given and consists of one historical series of building energy and $M-1$ additional energy-related feature series. Building energy prediction is mainly a time series regression problem, which is to learn a nonlinear mapping function [38] to obtain the predicted energy value Y. There could be a dataset$D = \{(X_1, Y_1), (X_2, Y_2), \cdots, (X_N, Y_N)\}$, as a collection of pairs(X_i, Y_i), where X_i could either be a univariate or multivariate series and usually would be preprocessed by the sliding window [39], and Y_i could be more than one value for multi-step building energy prediction [10].

The ANNs for building energy prediction are designed to learn hierarchical representations of energy data, which could be complex machine learning models when adding more layers, so-called deep learning [40]. A general ANN could be considered as a composition of L mapping functions, so-called layers, which is a representation of the input domain [41]. The mapping function is controlled by a set of parameters θ_l for each layer, so-called weights and bias. Given an input $\times$, the general computation would be performed in ANNs [42], so-called feedforward propagation, as the following:

$$\sigma_L(\theta_L, x) = \sigma_{L-1}(\theta_{L-1}, \sigma_{L-2}(\theta_{L-2}, \cdots, \sigma_1(\theta_1, x)) \quad (1)$$

where σ_j is the nonlinearity or activation function applied in layer j, and θ_j is the vector of parameters in layer j, such as weights and biases.

During the training phase, ANNs would be trained with a certain number of known input–output pairs. After the weights in ANNs are usually initialized randomly, a feedforward propagation would be implemented, and the prediction loss would be computed by a cost function, i.e. *mean absolute error* or *mean square error* for the regression problem [43,44]. Then, the weights would be updated in a backpropagation using gradient descent. The training phases iteratively take a feedforward propagation followed by a backpropagation to update the parameters of the ANNs to minimize the loss on the training data [45]. During the test phase, the ANNs would be tested on unseen data and compute the prediction error to measure the generalization ability.

In the following subsections, a detailed survey from the selected publications will be grouped by different ANN architectures applied in building energy prediction, whose taxonomy refers to [25,46], including FFNN, RNN, and CNN.

3.2. Feed forward neural networks

Fig. 4 illustrates the general architecture of FFNNs. MLP is the simplest and the most widely used type of FFNNs, which is also known as a fully-connected networks. The general formulation of MLP is shown as the following:

$$H_l = \sigma(w_l X + b_l) \quad (2)$$

where$w$ and b is the weights and bias respectively, H is the activation or output of the neurons, and σ is the activation function, which usually is the *sigmoid* function to squashes inputs to a [0,1] range in MLP [44,47] as the following:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (3)$$

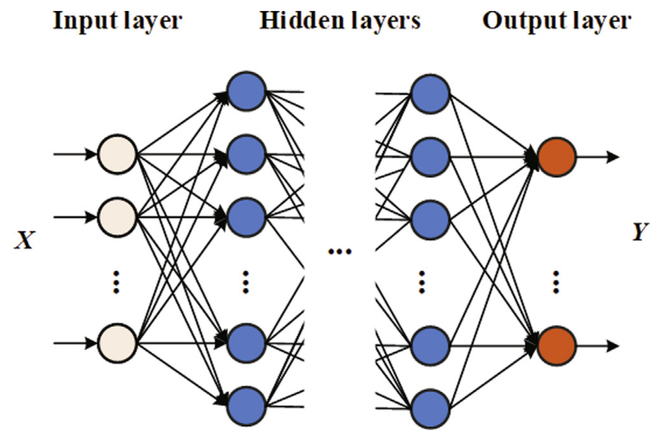


Fig. 4. The general architecture of FFNNs.

MLP often appears in building energy prediction as the bench model, and has certain advantages when compared to other machine learning models. Gunay et al. [48] found the a single hidden layer MLP could outperform better than MLR in characterizing the heating and cooling load pattern with one-hour history of weather and electricity load. Ahmad et al. [47] compared MLP with tree-based models and the results indicated MLP performed marginally better for hourly electricity consumption. Magalhaes et al. [43] concluded MLP could estimate effectively the heat energy use both at an individual and the buildings stock level. Seyedzadeh [49] found MLP would be more appropriate for building load prediction when the data was complex and abundant.

RBF is a type of FFNNs with single hidden layer which is used to map the input data to the network space using a radial basis function. The general formulation of RBF is as the following:

$$y = \sum_{j}^{N} w_j \varphi_j(\|x - v_j\|) + b \tag{4}$$

where v_j is the jth center location of the RBF nodes, $\|\cdot\|$ is the Euclidean distance and φ is the radial basis function which usually is the Gaussian function as the following:

$$\varphi_j(x) = \exp\left(-\frac{(x - v_j)^2}{2\sigma_j^2}\right) \tag{5}$$

In which σ is the width parameters of the RBF nodes. Due to the structure simplicity, the adaptation of two typical-stage training schemes and fewer control parameters, RBF has an easier and shorter training phase and has good generalization [50,51]. Tran et al. [52] developed an ensemble model based on the RBF and the least-square SVR to prediction energy consumption in residence buildings. They utilized an optimization technique, symbiotic organisms search, to automatically optimize the number of hidden neurons and the width of Gaussian function. The ensemble model was proven as the most effective model.

Considering the slow gradient learning and massive parameters to be learned of traditional FFNNs, ELM is proposed as a modification of a single hidden layer FFNN to provide good generalization performance at extremely fast learning speed, which solves the drawbacks caused by gradient descent algorithms [53]. The key principle behind ELM is that input weights and biases are randomly generated and should not be tuned in the training phase [53]. Thus, training the ELM could transfer to solve a linear system and the output weights could be calculated as the following:

$$\beta = \mathrm{H}^{+}Y \tag{6}$$

where H^{+} is the Moore-Penrose generalized inverse of the output matrix H of the hidden layer, which can be calculated when $\mathrm{H}^{+}H$ is non-singular and $\mathrm{H}^{+} = (\mathrm{H}^{T}H)^{-1}\mathrm{H}^{T}$. Guo et al. [54] found the performances of the ELM were better than multiple linear regression, the SVR and the MLP for predicting demand in the building heating systems. Sekhar et al. [55] tried a hybrid model of the multivariate adaptive regression splines and the ELM for heating load prediction, which boosted the performance and outperformed GPR, MLR, MLP and RBF. Song et al. [56] proposed a framework by a two-step evolutionary algorithm, called evolutionary model construction, for channel selection, feature extraction and prediction in accurate electricity consumption prediction. The implementation of the framework chose the particle swarm optimization (PSO) with the ELM and the random vector functional link and the superiority of the framework was proven by comparison with the existing approaches. Naji et al. [57] used the ELM and genetic programming to estimate building energy consumption with five different wall details with various layer thicknesses in the design phase. Considering the situation where all the load data is not available in buildings, Kumar et al. [58] applied online sequential ELM to learn in chunk by chunk from recent examples and they improved the feature selection by combining the PSO [59]. Fayaz et al. [60] used the deep ELM for energy consumption prediction in residential buildings and its performance was far better than ANFIS and MLP.

WNN is a type of FFNN combining the wavelet theory and ANNs. In WNN, the transfer function is the wavelet basis function and the output of the hidden layer is calculated as the following:

$$h = \phi\left(\frac{\sum_{i=1}^{k} w_i x - \beta}{\alpha}\right) \tag{7}$$

$$\phi(z) = cos(1.75z)e^{-z^2/2} \tag{8}$$

where ϕ is the wavelet basis function, α and β are the scaling factor and translation factor of the wavelet basis function, respectively. Gu et al. [61] found the WNN could obtain smaller relative error when predicting medium-term (one day ahead) heating load for an existing residential building. Yuan et al. [62] developed a hybrid load forecasting engine for commercial buildings, combining the singular spectrum analysis and WNN, and used the cuckoo search for parameter tuning of WNN.

3.3. Recurrent neural networks

Fig. 5 illustrate of the general architecture of RNNs. RNNs are often employed for modeling time-series data due to their strength in modeling dependencies across time. They have been successfully applied to temporal dependency extraction and complex sequence modeling, showing effectiveness in many research fields, such as neural language processing [63] and video action recognition [64]. The general formulation of RNN (Vanilla RNN) is as follows.

$$h_t = \sigma(W_x X + W_h h_{t-1} + b) \tag{7}$$

where h_t is the hidden state at time step t.

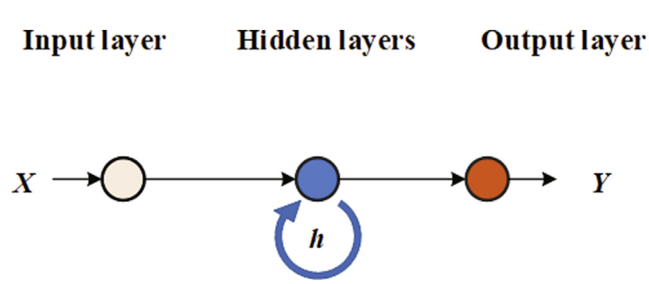


Fig. 5. The general architecture of RNNs.

The nonlinear autoregressive (NAR) and the nonlinear autoregressive neural network with exogenous inputs (NARX) could be considered as the basic types of RNNs, and the architecture of NARX is shown in Fig. 6. NAR makes final prediction only using the historical values, so-called time delays, and NARX includes another external series that may provide additional useful information. Ruiz et al. [65] used NAR and NARX for energy prediction in an university buildings. The result showed the prediction accuracy of NAR and NARX were both suitable. Due to adding the temperature as the exogenous input, NARX furnished a better performance, and decreased the network complexity and the amount of historical data. Kim et al. [66] also found NARX worked better than ARIMA and the exponential model for 1 h to 1 day ahead forecasting in an institutional building. Koschwitz et al. [67] predicted long-term urban heating load using NARX based on building retrofit effects.

Elman Neural Network (ENN) is another simple type of RNNs, which introduces the concept of memory [68]. In ENN, the output of the hidden layer in the previous time step is fed back as an input to the hidden layer in the current time step, shown in Fig. 7, which provides a better modelling of time-series with historical dependencies. The main difference between NARX and ENN is that there are the context units to store the state or output of hidden neurons. Ma et al. [69] employed an ENN to analyze the impacts of direct solar irradiance and wind speed on heat demand prediction in Finland. The results showed the ENN with a sliding window of 4 h and a number of layers of 8 could achieved the lowest MAPE of the predictions and Simultaneous involvement of both wind speed and direct solar irradiance performed the best. Bedi et al. [70] predicted the electric energy consumption along with the ambient temperature and the occupancy state, which showed the ENN model outperformed the exponential model on six out of seven days of electric energy consumption predictions. Ruiz et al. [71] used the genetic algorithm for optimizing the weight of the ENN to solve the disadvantages of slow convergence and local minimum stagnation, which showed an improvement in the accuracy of energy consumption forecasting for public buildings of University of Granada.

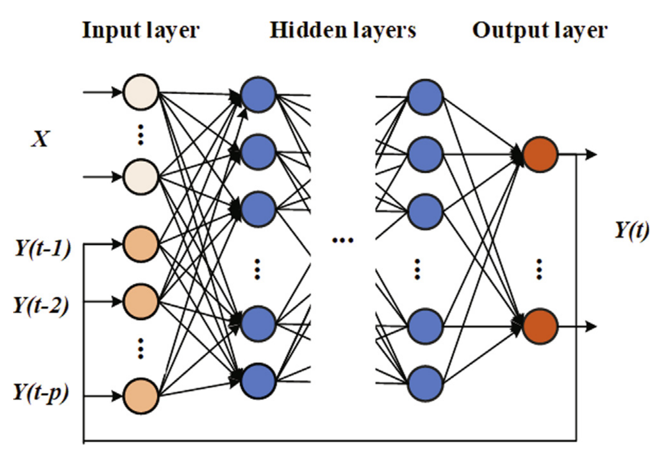


Fig. 6. The architecture of NARX.

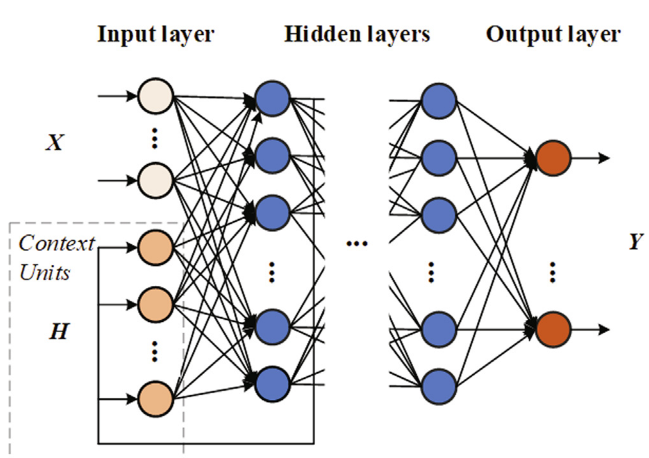


Fig. 7. The architecture of ENN.

LSTM is a special form of RNN, designed for handling long sequential data. The memory block of LSTM is shown in Fig. 8. LSTM contains three gates: the forget gate, the input gate, and the output gate denoted as *f*, *i*, and *o* respectively. The three gates could learn when to forget previous information and how to update them using new information [72], which helps solve the problem of a vanishing or exploding gradient when the number of time steps is large. Wang et al. [73] applied LSTM using occupancy and time index to predict the plug load, which was found to perform better than ARIMA approach. Moreover, they recommended LSTM for short-term (1 h ahead) cooling load prediction of a campus building with weather forecasting [74], which showed 20.2% in the coefficient of variation of the root mean square error (CVRMSE) and outperformed other statistical models and machine learning models. Somu et al. [39] applied the improved sine cosine optimization algorithm to identify the optimal hyperparameters of LSTM in real time for accurate and robust building energy consumption forecasting. Rahman et al. [75] found LSTM could be comparatively more accurate than MLP in predicting heating demand.

GRU is a popular variant of LSTM, which offers a simpler structure and might show better performance than the original LSTM in some cases. In GRU, the forget gate and input gate are replaced by a reset gate and an update gate, denoted as *r* and *z* respectively, whose memory block is shown in Fig. 9. Zhang et al. [76] utilized LSTM and GRU to extract features from raw load data in a public building, and utilized MLP to learn the relationships in the features extracted, which found GRU had greater potential than LSTM in building energy prediction. Wen et al. [77] used a deep GRU to investigate the load prediction of residence buildings over the short to medium-term, which achieved higher accuracy compared to LSTM and other conventional models.

The RNN architectures would also influence the performance of building energy prediction [78,79]. In general, there are three popular RNN architectures for time series prediction, including the direct architecture, the recursive architecture and the multi-input and multi-output architecture [80]. Considering the standard recurrent architectures are not naturally suited to model relation-

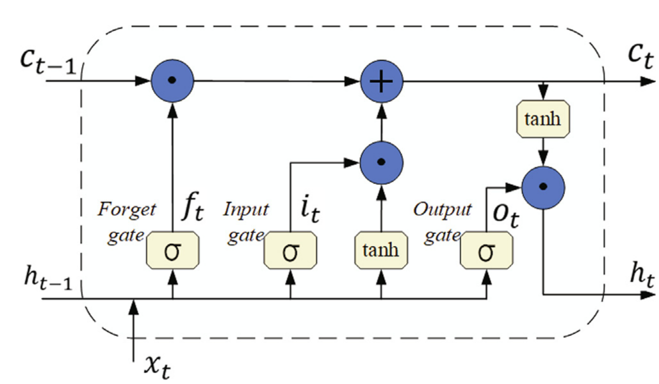


Fig. 8. The memory block of LSTM.

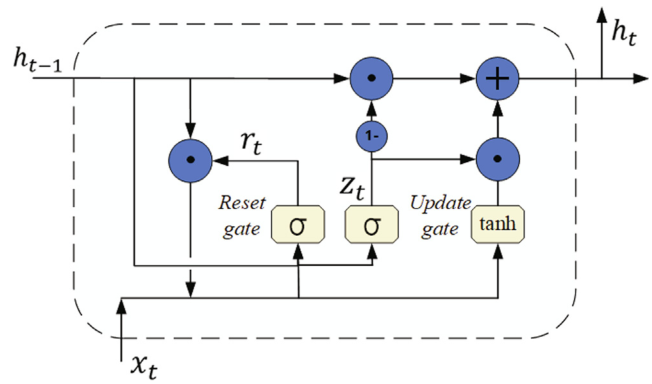


Fig. 9. The memory block of GRU.

ships between input and output sequences whose lengths are different, the sequence-to-sequence (Seq2Seq) model are used to address the problem. Skomski et al. [81] applied a Seq2Seq LSTM model to predict short-term electrical load in four office buildings and investigated the impacts of the time resolutions, the amount of training data available and the lengths of the input and the output sequences. The Seq2Seq LSTM model was effective. Furthermore, the transferability of the model across different buildings was considered and the results showed it was highly dependent on building pairs. Attention mechanism was design to help remember long-term input history and especially paid attention to those critical input. Chitalia et al. [39] combined LSTM and bidirectional LSTM (BiLSTM) with attention to predict five commercial buildings. Sehovac et al. [82] proposed a Seq2Seq RNN with two main different attention mechanisms for load prediction, which were named Bahdanau attention [83] and Luong attention [84] respectively. They found the Seq2Seq RNN with Bahdanau attention outperformed the other models.

Echo State Network (ESN) is another type of RNNs with powerful nonlinear time series modeling ability, which has the advantages of fast convergence and global optimal solution [85,86]. The ESN generally comprises an input layer, a dynamic reservoir and an output layer, shown in Fig. 10. The dynamic reservoir contains several sparsely connected neurons and the output layer is a memoryless linear readout trained to generate the output. Shi et al. [87] developed the ESN with several topologies to predict the energy consumption in different room types of an office building, showing the excellent performance. Mansoor et al. [88] compared the ESN and the FFNN, and the ESN performed slightly better than the FFNN in the zone-based analysis, but conversely in the cluster-based analysis. Wang et al. [89] improved the ESN for electricity consumption forecasting by using the differential evolution algorithm (ESN-DE) to search optimal value of three crucial parameters in the ESN, including the scale of the reservoir, the connectivity rate among the neurons of the reservoir and the spectral radius. The results showed the ESN-DE outperformed the traditional ESN. Hu et al. [90] developed a deep ESN with a stacked hierarchy of reservoirs, which outperformed the persistence model, the MLP and the traditional ESN for forecasting energy consumption and wind power generation. Then, they continued proposing an model combing ESN, bagging and differential evolution algorithm, and achieved better performance in accuracy and reliability [91].

Restricted Boltzmann Machine (RBM) is a generative stochastic bool ANN to learn a probability distribution over its set of inputs and is trained by minimizing a pre-defined energy function to learn an equilibrium from the visible layer to the hidden layer, whose architecture is shown in Fig. 11. The RBM could be considered as a type of RNN. Fu [92] established an ensemble model of empirical mode decomposition and deep belief network (DBN) to forecast cooling load, which could be view as a stack of several RBMs. The model exhibited competitive performance. Hafeez et al. [93] proposed a novel short-term electric load forecasting model based on factored conditional RBM using modified mutual information to preprocess data and genetic wind-driven algorithm to optimize. The factored conditional RBM had a rich, distributed hidden state to help in preserving the temporal information in the electric data, and it prevented the vanishing gradient issue in back propagation.

3.4. Convolutional neural networks

CNN is originally invented for image processing, but it also has been proven effective in the time-series problems [94]. CNN has two basic operations, convolution and pooling. The convolution could be seen a sliding filter over the energy series, which usually exhibits one-dimension, so-called 1D CNN. The pooling, such as average and max, could reduce the length of input series by aggregating over a sliding window [42]. Sadaei et al. [95] proposed a combined model of CNN and fuzzy time series for short-term load forecasting. The fuzzy time series was applied to convert the original input series to the format of images in the input layer of CNN. Cai et al. [96] developed a gated CNN for day-ahead building-level load forecasts. The gate CNN introduced a novel gating mechanism, where each input was processed simultaneously using two different convolutional operations for detecting the temporal dependency of the neighboring time stamps and identifying the weather correlation at each time stamp respectively. The results indicated the gate CNN outperformed other models from all the

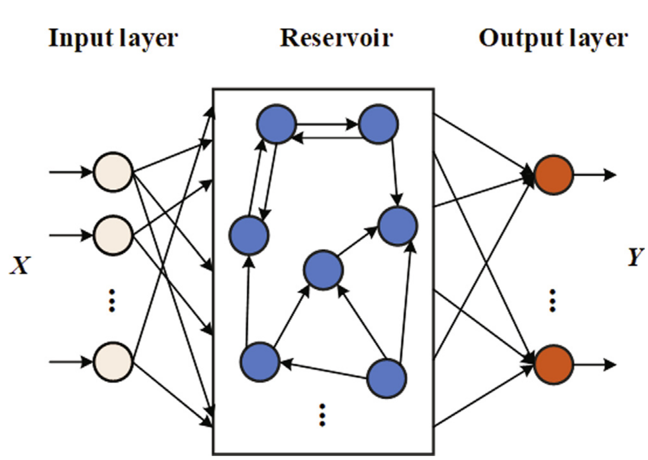


Fig. 10. The architecture of ESN.

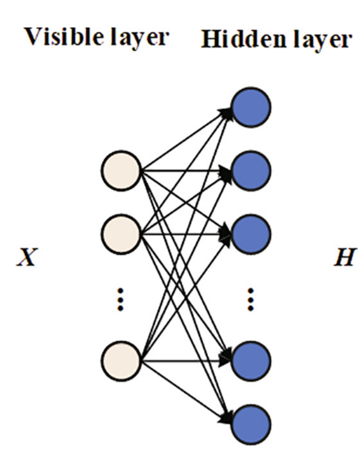


Fig. 11. The architecture of RBM.

aspects of prediction accuracy, computational efficiency and generalizability.

It is more common that CNN is generally as feature extraction layers for other machine learning models. The hybrid models of CNN and RNN are the most common combination for building energy prediction [97,79,97,98], where CNN could obtain the local and spatial features and the features would be fed into RNN to capture the temporal dependencies in the time series [99]. Chitalia et al. [39] used the CNN-LSTM model and CNN-BiLSTM model to predict electric load. They also used the ConvLSTM model and ConvBiLSTM model that were when gates performed convolutions in LSTM and BiLSTM respectively [100]. Sajjad et al. [98] developed a hybrid model incorporating CNN with GRU and achieved better performance in terms of preciseness and efficiency. Another popular way to combine CNN and RNN is based on the inception module [101]. Kim et al. [67] proposed a recurrent inception CNN for short-term load forecasting that combined LSTM and one-dimensional CNN, where the CNN could calibrate the hidden state vector values generated from LSTM. The result showed the proposed model could yield better forecasting than MLP, CNN and RNN.

4. How to choose suitable ANN architectures for building energy prediction

In the last section, we have illustrated the ability and applications of the twelve ANN architectures in the literature. However, the various implementation and different performance results exist in those literature so that choosing the suitable ANN architecture for better building energy prediction is still an open issue for researchers. In this section, we will provide some overall statistics and summarize some notable observations about the current state and research trend about building energy prediction using ANNs, which is meant to recommend ANN architectures in practice for new researchers.

The research trends of different ANN architectures in building energy prediction could be explored in Fig. 12, which shows the number of journal articles based on different ANN architectures in building energy prediction. First, the number of journal articles based on FFNNs accounts for the majority of every year and still increases. Second, the proportion of journal articles based on RNNs has rapidly increased since 2017, which is 17.39% and 40.4% in 2017 and 2020 respectively. Third, the number of journal articles where CNN is dominant in structures are rare and only 2 and 1 publications in 2019 and 2020 respectively.

Overall, FFNNs still is the most popular choice in building energy prediction because they are easy to implement. But FFNNs don't have the ability to sequence modeling, so they process the relationship between energy-related features and building energy data in most studies. Careful feature engineering is usually required to determine the inputs before FFNNs model. Taking historical energy data as inputs is proven to improve prediction performance. The lag and number of historical energy data also need to be determined by feature engineering.

Most building energy data includes time-dependent components, so naturally, RNNs have been becoming another preferred choice of most researchers for building energy prediction. Different from FFNNs, RNNs usually can process directly the building energy data and learn the historical information by themselves. Fig. 13 provides the distribution of RNNs in our selected studies. LSTM has attracted the most attention among all RNNs with 42 journal articles (account for 48.93%). It means LSTM is dominant in ANN structures or LSTM outperforms other models in these 42 articles. NAR/NARX, as the basic RNN model, has 13 journal articles (account for 15.15%). Hence, it is clear that LSTM is the most popular RNN model and should be recommended at least for benchmarking in building energy prediction currently.

CNN is also a potential choice for building energy prediction. There exists 1D CNN proposed to process time series, including building energy prediction. A common and effective practice is recommended to combine CNN and RNN (CNN-RNN). In CNN-RNN, convolution and pooling layers were used to capture effective features and reduce the problem dimensionally while greatly suppressing the redundancy in representations of fine-grained building energy data. RNN can capture temporal dependency from the output of CNN. At least, CNN-RNN is also worthy as another baseline model for comparison. It is worth mentioning that CNN is widely applied for image recognition applications. Hence, it should be more attempted to convert the building energy data into an image-like two-dimensional representation in an innovative way as the input of CNN.

How to choose a suitable ANN architecture for a specific building energy prediction task is still an open issue. One of the future directions is to attempt more advanced ANN architectures for better building energy prediction. Customization of ANN architectures could always be effective for improving prediction performance. Furthermore, researchers should explore and summarize whether there exists the general or suitable ANN architecture when dealing with different building energy prediction tasks, which may involve the interpretability of ANN.

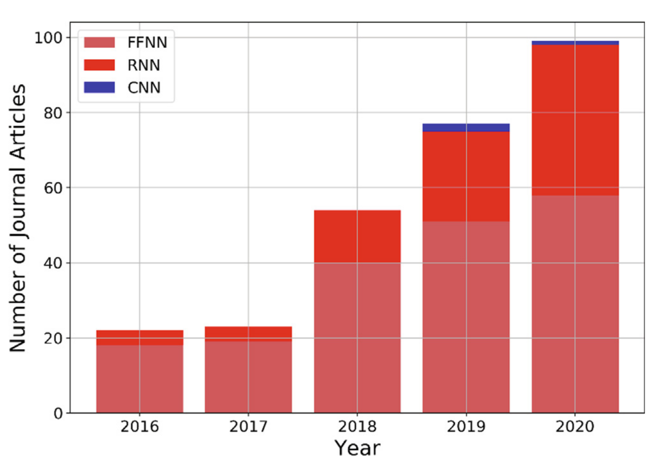


Fig. 12. The number of journal articles based on different ANN architectures in building energy prediction per year.

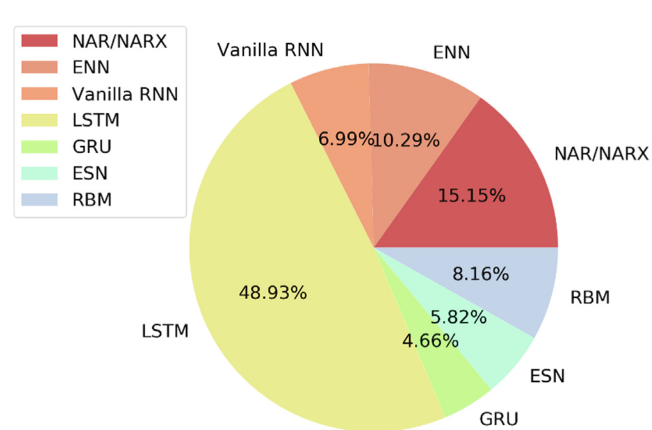


Fig. 13. The distribution of journal articles based on different RNNs.

5. How to improve performance of ANNs for building energy prediction

In this section, we will introduce how to improve performance of building energy prediction with precision and stability when implementing ANNs in the existing studies.

Implementing ANNs in building energy prediction always encounter three main problems. The first is that ANNs initialize randomly and easily result in the local optimal problem and the unstable performance during the training phase [102], so weights and biases need to be updated more reasonably. The second is the input selection problem because it is hard to decide inputs for ANNs [56,103], such as the number of energy-related features and the length of historical energy values. The third is the hyperparameters setting problem, which is a tedious and time-consuming task in deep learning especially. The hyperparameters will impact the prediction performance and need to be optimally determined [103,104], such as the number of layers, neurons, and epochs, and the type of optimizers and activation functions. To overcome these problems, three types of parameters need to be optimized for ANNs in building energy prediction, including weights and biases, input features and hyperparameters respectively.

In our selected studies, *meta*-heuristic algorithms are widely utilized and proven effective solutions for those problems, improving the reliability of ANNs and helping achieve the best prediction performance. Meta-heuristic algorithms could try to evading many local minima or conduct global search to decide the parameters. A summary of the optimization algorithms for ANNs in building energy prediction is shown in Table 3. Most *meta*-heuristic algorithms are inspired by the observation of natural behaviors, which could be classified into four categories [114], including evolutionary-based, physical or mathematic-based, human-based and swarm intelligence-based. The evolutionary-based algorithms are inspired by the biological evolutionary, including evolutionary algorithms (EA), differential evolution (DE), and genetic algorithm (GA). The human-based algorithms are inspired from behaviors of human beings, including sine cosine optimization algorithm (SCOA), imperialist competitive algorithm (ICA) and teaching–learning-based optimization (TLBO). The swarm intelligence-based algorithms are inspired by the communities in herds of animals, colonies of insects and flocks of birds, including particle swarm optimization (PSO) algorithm, bat algorithm (BA), artificial bee colony (ABC) algorithm, cuckoo search (CS), symbiotic organism search (SOS), fruit fly optimization algorithm (FFOA), and dragonfly algorithm (DA). We summarize briefly the advantages and disadvantages of the mentioned optimization algorithms according to our selected studies in Table 4. Regarding optimization algorithms, more details and options can refer to [114–116]. For the interested researchers, it is important to know that every algorithm is essentially designed to expand the search space to obtain an optimal solution for different parameters of ANNs. In terms of prediction performance and computational time, there does not exist a general algorithm for all types of real-world problems, which motivates researchers to modify the existing algorithms or develop new algorithms for their own problems and needs.

Table 3
The summary of the optimization algorithms for ANNs in building energy prediction.

Reference	ANN architectures	Optimization Algorithms	Parameters to be optimized
[105]	MLP	GA	Input features
[61]	MLP	GA	Weights and biases
[17]	MLP	GA, ICA	Input features, Hyperparameters
[102]	MLP	ABC, GA, PSO, ICA	Weights and biases
[106]	MLP	GA	Hyperparameters
[107]	MLP	ABC, PSO	Weights and biases
[108]	MLP	EA	Hyperparameters
[109]	MLP	TLBO, CS	Weights and biases
[52]	RBF	SOS	Weights and biases
[56]	ELM	PSO	Input features
[110]	ELM	Bat algorithm	Hyperparameters
[62]	WNN	CS	Weights and biases
[71]	ENN	GA	Weights and biases
[111]	ENN	DA	Weights and biases
[39]	LSTM	SCOA	Hyperparameters
[103]	LSTM	GA, PSO	Input features, Hyperparameters
[104]	LSTM	GA	Hyperparameters
[89]	ESN	DE	Hyperparameters
[112]	ESN	FFOA	Input features

Another approach to improve the predictive performance that should be mentioned is online learning strategies. Most of our selected studies belong to the conventional offline learning, which belongs to the batch learning strategy using a static data set and if dealing with the new data, ANNs need to be re-trained. If the pattern changes in the new energy data, ANNs using offline learning would cause degrading predictive performance. Building energy data usually is generated as a data stream in practice, and online learning strategies would dynamically adapt to new patterns in the data stream and adopt different architecture and parameters. The data could be discarded after they are learned. Fekri et al. [115] proposed online adaptive RNN, which adopted online preprocessing techniques to prepare data for the RNN model and adjusted hyperparameters (learning rate) when the predictive performance started to deteriorate. The online adaptive RNN not only outperformed the other online model and the offline RNN but also reduced the training time significantly compared to the offline RNN.

The future research directions include developing hybrid models to obtain better performance, which should combine ANNs and suitable optimization algorithms, and implementing the online learning strategies in practice. The hybrid models would either achieve the outperforming results or avoid massive manual parameters tuning of complex ANNs by trial-and-error effectively. The online learning strategies would reduce the computational time and adapt to the new pattern in the energy data stream.

6. How to deal with the lack of energy training data

In this section, we will discuss another challenge in building energy prediction, how to deal with the lack of energy training data when implementing ANNs. Although lots of advanced and complex ANN architectures have been successfully applied and achieved more and more precise prediction, the ANNs rely on a large amount of historical energy data as the training data. The challenge is that most buildings cannot provide sufficient energy data and data collection is time-consuming, especially for the newly built buildings and buildings without IoT infrastructures.

6.1. Transfer learning

Transfer learning focuses on transferring the knowledge across different domains, which is considered as a promising machine learning methodology for solving the challenge above [116,117]. The core of transfer learning is to find the similarity between source and target domains, and improve the model learning in the target domain through using the knowledge learned from the source domain. In building energy prediction, the source domains refer to the buildings with abundant energy data. The target domains refer to the buildings lacking of energy data, where cannot well trained the ANNs. There exist a few studies showing the beneficial of transfer learning [118–122]. However, the researches

Table 4
The advantages and disadvantages of the optimization algorithms for ANNs.

Categories	Optimization Algorithms	Inspiration	Advantages	Disadvantages
Evolutionary-based	EA	The natural mechanisms of the genetic evolution of biological species.	Easy to implement; Suitable for various solutions.	Computational time might be long; Large amounts of computing resources for difficult problems.
	DE	The basic rules of genetics (mutation and crossover strategies).	Simple and efficient heuristic for global optimization; Few numbers of control parameters	Parameter tuning mostly by trail-and-error.
	GA	The natural process of fruition (selection, mutation and crossover).	Global search; Without prior knowledge (not depend on the initial solution)	Slow convergence speed; Can be complex when solving high dimensional problems.
Physical or mathematic-based	SCOA	Simple sine and cosine mathematical functions.	Minimal tuning parameters; High optimization accuracy; Fast convergence speed; Strong global search ability.	Poor local search ability.
Human-based	ICA	Imperialistic competition and based on a social policy of imperialism.	Find the global optimum solution; Few parameters to adjust; Easy to implement; Fast convergent.	Rapidly declining diversity; Premature convergence.
	TLBO	The influence of a teacher on the learning outcome of its students.	Balanced modeling accuracy and network complexity;	Easy to be trapped locally; Poor global search ability; Random search.
Swarm intelligence-based	PSO	The flocking of birds where each particles keeps personal and global best value and changed the velocity with position in each step.	Easy to implement; fewer tuning parameters; Robust.	Might stuck in local minima; Premature convergence for certain complex problems; low accuracy
	ABC	The foraging behavior of a swarm of bees, which comprises employed bees, onlookers, and scout bees.	Easy to implement; Robust against initialization.	Might stuck in local optimum; Poor exploitation characteristics;
	CS	The brood parasitic behavior of some cuckoo species and levy flights of some birds and flies.	High convergence speed; global search ability.	Slow convergencein late evolution; Easy to fall into local minimum.
	SOS	The interactions of organisms in nature, and involves mutualism, commensalism, and parasitism.	Easy adjustability of the common parameters; Simplicity of operation	Low optimization accuracy; Slow convergence in late period.
	FFOA	The food finding behavior of the fruit fly.	Low complexity, Fast computation speed; Universal solving ability	Low optimization accuracy; Might fall into local optimum.
	BA	The echolocation behavior of bats to search directions and location.	Random optimization; New features of echolocation.	Might fall into local optimum; Slow solution.
	DA	The static and dynamic behavior of dragonfly swarms	Converge towards the global optimum.	Might fall into local optimum; Slow convergence

using transfer learning in building energy prediction are at the primary stage and most of them just utilized the fine-tune with the deep learning models. One of the future directions is to explore more possibility to implement transfer learning, such as instance-based [123], feature-based [124] and relation-based [125] transfer learning, which may need investigate the inner distribution and relations in building energy data.

6.2. Generative learning

Generative learning is another solution to solve the problem of the lack of data. The representative of generative learning is generative adversarial network (GAN), which was first proposed in [126]. GAN consists of two networks, the generator and the discriminator. The generator usually generates noise data randomly and the discriminator judges whether the data is real or fake, so GAN could learn the data distribution and generate synthetic data with the same distribution. In building energy prediction, different from transfer learning, GAN leverage knowledge from the similar functional buildings to enhance data variety of the target building [127,128,129], which essentially is data augmentation. However, GAN has shown more applications in sequence prediction in [130,131,132]. Thus, one of the future directions is explore more applications of GAN in building energy prediction.

7. Conclusions

This paper presents a comprehensive literature survey on building energy prediction using ANNs in the past five years and a total of 324 related studies were selected. This paper summarizes twelve ANN architectures applied in those studies and introduced them in detail. Then, the current state and research trend of building energy prediction using ANNs are provided. LSTM is the most popular RNN and CNN-RNN is proven as an effective architecture, both of which are recommended as the benchmark for following researchers. Furthermore, the optimization algorithms are summarized for prediction performance improvement, which is usually applied to optimize three types of parameters of ANNs, including weights and biases, input features, and hyperparameters. Online learning should be implemented to adaptively optimize ANNs when dealing with the building energy data stream in practice. Moreover, there would exist a lack of energy data when buildings are newly completed or renovated, transfer learning and generative learning could be considered promising solutions.

Overall, based on the literature survey and analysis, some of the future research directions could be given below for the following researchers (Not limited to these):

- Customization of ANN architectures for one specific building energy prediction case;

- Combination of ANNs and suitable optimization algorithms to enhance building energy prediction performance;
- Implementation of online learning and efficient computing techniques for building energy prediction in practice;
- Exploration of the transferability of the different situations of buildings, such as different building functions and climate zones;
- Exploration of the availability of generative learning in different buildings.

CRediT authorship contribution statement

Chujie Lu: Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Data curation, Writing – original draft, Writing – review & editing. **Sihui Li:** Writing – review & editing. **Zhengjun Lu:** Writing – review & editing.

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.

Acknowledgements

This work was supported by the Leading Talents of Guangdong Province Program under Grant 2016LJ06D557. I would like to appreciate particularly Prof. *Jianguo Ma*, my Ph.D. supervisor, for the invaluable counsel and knowledgeable instruction in the research methodology. I also would like to appreciate Prof. *Qijun Zhang* at Carleton University, Canada, for his expert guidance in artificial neural networks. A special thanks goes to *Wenhao Sun* for all of his motivation.

References

[1] International Energy Agency, Tracking Buildings 2020, Paris, 2020.
[2] C.V. Gallagher, K. Bruton, K. Leahy, D.T.J. O'Sullivan, The suitability of machine learning to minimise uncertainty in the measurement and verification of energy savings, Energy Build. (2018), https://doi.org/10.1016/j.enbuild.2017.10.041.
[3] R. Jia, B. Jin, M. Jin, Y. Zhou, I.C. Konstantakopoulos, H. Zou, J. Kim, D. Li, W. Gu, R. Arghandeh, P. Nuzzo, S. Schiavon, A.L. Sangiovanni-Vincentelli, C.J. Spanos, Design Automation for smart building systems, Proc. IEEE 106 (2018) 1680–1699, https://doi.org/10.1109/JPROC.2018.2856932.
[4] B. Gunay, W. Shen, Connected and distributed sensing in buildings: improving operation and maintenance, IEEE Syst. Man Cybern. Mag. 3 (2017) 27–34, https://doi.org/10.1109/MSMC.2017.2702386.
[5] A. Kusiak, M. Li, Cooling output optimization of an air handling unit, Appl. Energy (2010), https://doi.org/10.1016/j.apenergy.2009.06.010.
[6] N. Luo, T. Hong, H. Li, R. Jia, W. Weng, Data analytics and optimization of an ice-based energy storage system for commercial buildings, Appl. Energy (2017), https://doi.org/10.1016/j.apenergy.2017.07.048.
[7] S. Li, J. Peng, Y. Tan, T. Ma, X. Li, B. Hao, Study of the application potential of photovoltaic direct-driven air conditioners in different climate zones, Energy Build. 226 (2020), https://doi.org/10.1016/j.enbuild.2020.110387 110387.
[8] X. Liu, A. Heller, P.S. Nielsen, CITIESData: a smart city data management framework, Knowl. Inf. Syst. (2017), https://doi.org/10.1007/s10115-017-1051-3.
[9] L. Zhang, J. Wen, Y. Li, J. Chen, Y. Ye, Y. Fu, W. Livingood, A review of machine learning in building load prediction, Appl. Energy 285 (2021), https://doi.org/10.1016/j.apenergy.2021.116452 116452.
[10] Y. Sun, F. Haghighat, B.C.M. Fung, A review of the-state-of-the-art in data-driven approaches for building energy prediction, Energy Build. 221 (2020), https://doi.org/10.1016/j.enbuild.2020.110022 110022.
[11] S. Zhang, Y. Chen, W. Zhang, R. Feng, A novel ensemble deep learning model with dynamic error correction and multi-objective ensemble pruning for time series forecasting, Inf. Sci. 544 (2021) 427–445, https://doi.org/10.1016/j.ins.2020.08.053.
[12] S. Panigrahi, H.S. Behera, A hybrid ETS–ANN model for time series forecasting, Eng. Appl. Artif. Intell. 66 (2017) 49–59, https://doi.org/10.1016/j.engappai.2017.07.007.
[13] S.D. Atalay, G. Calis, G. Kus, M. Kuru, Performance analyses of statistical approaches for modeling electricity consumption of a commercial building in France, Energy Build. 195 (2019) 82–92, https://doi.org/10.1016/j.enbuild.2019.04.035.
[14] H. Shakouri, G.A. Kazemi, Selection of the best ARMAX model for forecasting energy demand: case study of the residential and commercial sectors in Iran, Energ. Effi. (2016), https://doi.org/10.1007/s12053-015-9368-9.
[15] W. Gao, J. Alsarraf, H. Moayedi, A. Shahsavar, H. Nguyen, Comprehensive preference learning and feature validity for designing energy-efficient residential buildings using machine learning paradigms, Appl. Soft Comput. J. 84 (2019), https://doi.org/10.1016/j.asoc.2019.105748 105748.
[16] L. Alfieri, P. De Falco, Wavelet-based decompositions in probabilistic load forecasting, IEEE Trans. Smart Grid 11 (2020) 1367–1376, https://doi.org/10.1109/TSG.2019.2937072.
[17] A.T. Eseye, M. Lehtonen, Short-term forecasting of heat demand of buildings for efficient and optimal energy management based on integrated machine learning models, IEEE Trans. Ind. Inf. 16 (2020) 7743–7755, https://doi.org/10.1109/TII.2020.2970165.
[18] M. Shao, X. Wang, Z. Bu, X. Chen, Y. Wang, Prediction of energy consumption in hotel buildings via support vector machines, Sustainable Cities and Society. 57 (2020), https://doi.org/10.1016/j.scs.2020.102128 102128.
[19] M. Shen, Y. Lu, K.H. Wei, Q. Cui, Prediction of household electricity consumption and effectiveness of concerted intervention strategies based on occupant behaviour and personality traits, Renew. Sustain. Energy Rev. 127 (2020), https://doi.org/10.1016/j.rser.2020.109839 109839.
[20] L. Cai, J. Gu, Z. Jin, Two-Layer Transfer-Learning-Based Architecture for Short-Term Load Forecasting, IEEE Trans. Ind. Inf. 16 (2020) 1722–1732, https://doi.org/10.1109/TII.2019.2924326.
[21] E. Kamel, S. Sheikh, X. Huang, Data-driven predictive models for residential building energy use based on the segregation of heating and cooling days, Energy. 206 (2020), https://doi.org/10.1016/j.energy.2020.118045.
[22] S. Fathi, R. Srinivasan, A. Fenner, S. Fathi, Machine learning applications in urban building energy performance forecasting: A systematic review, Renew. Sustain. Energy Rev. 133 (2020), https://doi.org/10.1016/j.rser.2020.110287 110287.
[23] X.J. Luo, L.O. Oyedele, A.O. Ajayi, O.O. Akinade, Comparative study of machine learning-based multi-objective prediction framework for multiple building energy loads, Sustainable Cities and Society. 61 (2020), https://doi.org/10.1016/j.scs.2020.102283 102283.
[24] A. Kusiak, M. Li, Z. Zhang, A data-driven approach for steam load prediction in buildings, Appl. Energy 87 (2010) 925–933, https://doi.org/10.1016/j.apenergy.2009.09.004.
[25] L.C. Jain, M. Seera, C.P. Lim, P. Balasubramaniam, A review of online learning in supervised neural networks, Neural Comput. Appl. 25 (2014) 491–509, https://doi.org/10.1007/s00521-013-1534-4.
[26] S. Suresh, N. Sundararajan, R. Savitha, Erratum: Supervised Learning with Complex-valued Neural Networks, in (2013), https://doi.org/10.1007/978-3-642-29491-4_9.
[27] Y. Himeur, A. Alsalemi, F. Bensaali, A. Amira, Building power consumption datasets: Survey, taxonomy and future directions, Energy Build. 227 (2020), https://doi.org/10.1016/j.enbuild.2020.110404 110404.
[28] C. Deb, F. Zhang, J. Yang, S. Eang, K. Wei, A review on time series forecasting techniques for building energy consumption, Renew. Sustain. Energy Rev. 74 (2017) 902–924, https://doi.org/10.1016/j.rser.2017.02.085.
[29] Z. Wang, R.S. Srinivasan, A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models, Renew. Sustain. Energy Rev. 75 (2017) 796–808, https://doi.org/10.1016/j.rser.2016.10.079.
[30] K. Amasyali, N.M. El-gohary, A review of data-driven building energy consumption prediction studies, Renew. Sustain. Energy Rev. 81 (2018) 1192–1205, https://doi.org/10.1016/j.rser.2017.04.095.
[31] M. Bourdeau, X. Qiang Zhai, E. Nefzaoui, X. Guo, P. Chatellier, Modeling and forecasting building energy consumption: A review of data-driven techniques, Sustainable Cities and Society. 48 (2019), https://doi.org/10.1016/j.scs.2019.101533 101533.
[32] Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, X. Zhao, A review of data-driven approaches for prediction and classification of building energy consumption, Renew. Sustain. Energy Rev. 82 (2018) 1027–1047, https://doi.org/10.1016/j.rser.2017.09.108.
[33] Y. Wang, Q. Chen, T. Hong, C. Kang, Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges, IEEE Trans. Smart Grid (2019), https://doi.org/10.1109/TSG.2018.2818167.
[34] N.D. Roman, F. Bre, V.D. Fachinotti, R. Lamberts, Application and characterization of metamodels based on artificial neural networks for building performance simulation: A systematic review, Energy Build. 217 (2020), https://doi.org/10.1016/j.enbuild.2020.109972 109972.
[35] S.R. Mohandes, X. Zhang, A. Mahdiyar, A comprehensive review on the application of artificial neural networks in building energy analysis, Neurocomputing. 340 (2019) 55–75, https://doi.org/10.1016/j.neucom.2019.02.040.
[36] Y. Wang, Q. Chen, T. Hong, C. Kang, Review of smart meter data analytics: Applications, methodologies, and challenges, IEEE Trans. Smart Grid 10 (2019) 3125–3148.
[37] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, D. Altman, G. Antes, D. Atkins, V. Barbour, N. Barrowman, J.A. Berlin, J. Clark, M. Clarke, D. Cook, R. D'Amico, J.J. Deeks, P.J. Devereaux, K. Dickersin, M. Egger, E. Ernst, P.C. Gøtzsche, J. Grimshaw, G. Guyatt, J. Higgins, J.P.A. Ioannidis, J. Kleijnen, T. Lang, N. Magrini, D. McNamee, L. Moja, C. Mulrow, M. Napoli, A. Oxman, B. Pham, D.

Rennie, M. Sampson, K.F. Schulz, P.G. Shekelle, D. Tovey, P. Tugwell, Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement, PLoS Med. (2009), https://doi.org/10.1371/journal.pmed.1000097.
[38] Y. Li, Z. Zhu, D. Kong, H. Han, Y. Zhao, Knowledge-Based Systems EA-LSTM : Evolutionary attention-based LSTM for time series prediction, Knowl.-Based Syst. 181 (2019), https://doi.org/10.1016/j.knosys.2019.05.028 104785.
[39] N. Somu, G.R. M R, K. Ramamritham, A hybrid model for building energy consumption forecasting using long short term memory networks, Appl. Energy. 261 (2020), https://doi.org/10.1016/j.apenergy.2019.114131.
[40] Y. Lecun, Y. Bengio, G. Hinton, Deep learning, Nature (2015), https://doi.org/10.1038/nature14539.
[41] N. Papernot, P. McDaniel, Deep k-nearest neighbors: Towards confident, interpretable and robust deep learning, ArXiv. (2018).
[42] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, P.A. Muller, Deep learning for time series classification: a review, Data Min. Knowl. Disc. 33 (2019) 917–963, https://doi.org/10.1007/s10618-019-00619-1.
[43] S.M.C. Magalhães, V.M.S. Leal, I.M. Horta, Modelling the relationship between heating energy use and indoor temperatures in residential buildings through Artificial Neural Networks considering occupant behavior, Energy Build. 151 (2017) 332–343, https://doi.org/10.1016/j.enbuild.2017.06.076.
[44] J.S. Chou, D.K. Bui, Modeling heating and cooling loads by artificial intelligence for energy-efficient building design, Energy Build. 82 (2014) 437–446, https://doi.org/10.1016/j.enbuild.2014.07.036.
[45] Y.A. LeCun, L. Bottou, G.B. Orr, K.-R. Müller, Efficient BackProp BT - Neural Networks: Tricks of the Trade, in, Neural Networks: Tricks of the Trade (2012).
[46] H. Moayedi, M. Mosallanezhad, A.S.A. Rashid, W.A.W. Jusoh, M.A. Muazu, A systematic review and meta-analysis of artificial neural network application in geotechnical engineering: theory and applications, Neural Comput. Appl. 32 (2020) 495–518, https://doi.org/10.1007/s00521-019-04109-9.
[47] M.W. Ahmad, M. Mourshed, Y. Rezgui, Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption, Energy Build. 147 (2017) 77–89, https://doi.org/10.1016/j.enbuild.2017.04.038.
[48] B. Gunay, W. Shen, G. Newsham, Inverse blackbox modeling of the heating and cooling load in office buildings, Energy Build. 142 (2017) 200–210, https://doi.org/10.1016/j.enbuild.2017.02.064.
[49] S. Seyedzadeh, F. Pour Rahimian, P. Rastogi, I. Glesk, Tuning machine learning models for prediction of building energy loads, Sustain. Cities Soc. 47 (2019), https://doi.org/10.1016/j.scs.2019.101484 101484.
[50] T. Jain, S.N. Singh, S.C. Srivastava, Fast static available transfer capability determination using radial basis function neural network, Appl. Soft Comput. J. (2011), https://doi.org/10.1016/j.asoc.2010.11.006.
[51] M.Y. Cheng, M.T. Cao, D.H. Tran, A hybrid fuzzy inference model based on RBFNN and artificial bee colony for predicting the uplift capacity of suction caissons, Autom. Constr. (2014), https://doi.org/10.1016/j.autcon.2014.02.008.
[52] D.H. Tran, D.L. Luong, J.S. Chou, Nature-inspired metaheuristic ensemble model for forecasting energy consumption in residential buildings, Energy. 191 (2020), https://doi.org/10.1016/j.energy.2019.116552 116552.
[53] G. Bin Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: Theory and applications, Neurocomputing. (2006), https://doi.org/10.1016/j.neucom.2005.12.126.
[54] Y. Guo, J. Wang, H. Chen, G. Li, J. Liu, C. Xu, R. Huang, Y. Huang, Machine learning-based thermal response time ahead energy demand prediction for building heating systems, Appl. Energy 221 (2018) 16–27, https://doi.org/10.1016/j.apenergy.2018.03.125.
[55] S. Sekhar Roy, R. Roy, V.E. Balas, Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of MARS and ELM, Renew. Sustain. Energy Rev. 82 (2018) 4256–4268, https://doi.org/10.1016/j.rser.2017.05.249.
[56] H. Song, A.K. Qin, F.D. Salim, Evolutionary model construction for electricity consumption prediction, Neural Comput. Appl. 32 (2020) 12155–12172, https://doi.org/10.1007/s00521-019-04310-w.
[57] 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, https://doi.org/10.1016/j.energy.2015.11.037.
[58] S. Kumar, S.K. Pal, R.P. Singh, A novel method based on extreme learning machine to predict heating and cooling load through design and structural attributes, Energy Build. 176 (2018) 275–286, https://doi.org/10.1016/j.enbuild.2018.06.056.
[59] S. Kumar, S.K. Pal, R. Singh, A novel hybrid model based on particle swarm optimisation and extreme learning machine for short-term temperature prediction using ambient sensors, Sustainable Cities and Society. 49 (2019), https://doi.org/10.1016/j.scs.2019.101601 101601.
[60] M. Fayaz, D. Kim, A prediction methodology of energy consumption based on deep extreme learning machine and comparative analysis in residential buildings, Electronics (Switzerland). 7 (2018), https://doi.org/10.3390/electronics7100222.
[61] J. Gu, J. Wang, C. Qi, C. Min, B. Sundén, Medium-term heat load prediction for an existing residential building based on a wireless on-off control system, Energy. 152 (2018) 709–718, https://doi.org/10.1016/j.energy.2018.03.179.
[62] Z. Yuan, W. Wang, H. Wang, S. Mizzi, Combination of cuckoo search and wavelet neural network for midterm building energy forecast, Energy. 202 (2020), https://doi.org/10.1016/j.energy.2020.117728 117728.
[63] T. Young, D. Hazarika, S. Poria, E. Cambria, Recent trends in deep learning based natural language processing [Review Article], IEEE Comput. Intell. Mag. (2018), https://doi.org/10.1109/MCI.2018.2840738.
[64] H. Cheng, L. Yang, Z. Liu, Survey on 3D hand gesture recognition, IEEE Trans. Circuits Syst. Video Technol. (2016), https://doi.org/10.1109/TCSVT.2015.2469551.
[65] L. Ruiz, M. Cuéllar, M. Calvo-Flores, M. Jiménez, An application of non-linear autoregressive neural networks to predict energy consumption in public buildings, Energies. 9 (2016) 684, https://doi.org/10.3390/en9090684.
[66] Y. Kim, H. Gu Son, S. Kim, Short term electricity load forecasting for institutional buildings, Energy Rep. 5 (2019) 1270–1280, https://doi.org/10.1016/j.egyr.2019.08.086.
[67] J. Kim, J. Moon, E. Hwang, P. Kang, Recurrent inception convolution neural network for multi short-term load forecasting, Energy Build. 194 (2020) 328–341, https://doi.org/10.1016/j.enbuild.2019.04.034.
[68] J.L. Elman, Finding structure in time, Cogn. Sci. (1990), https://doi.org/10.1016/0364-0213(90)90002-E.
[69] Z. Ma, J. Xie, H. Li, Q. Sun, F. Wallin, Z. Si, J. Guo, Deep neural network-based impacts analysis of multimodal factors on heat demand prediction, IEEE Trans. Big Data 6 (2019) 594–605, https://doi.org/10.1109/tbdata.2019.2907127.
[70] G. Bedi, G.K. Venayagamoorthy, R. Singh, Development of an IoT-driven building environment for prediction of electric energy consumption, IEEE Internet Things J. 7 (2020) 4912–4921, https://doi.org/10.1109/JIOT.2020.2975847.
[71] L.G.B. Ruiz, R. Rueda, M.P. Cuéllar, M.C. Pegalajar, Energy consumption forecasting based on Elman neural networks with evolutive optimization, Expert Syst. Appl. 92 (2018) 380–389, https://doi.org/10.1016/j.eswa.2017.09.059.
[72] G. Van Houdt, C. Mosquera, G. Nápoles, A review on the long short-term memory model, Artif. Intell. Rev. (2020), https://doi.org/10.1007/s10462-020-09838-1.
[73] Z. Wang, T. Hong, M.A. Piette, Predicting plug loads with occupant count data through a deep learning approach, Energy. 181 (2019) 29–42, https://doi.org/10.1016/j.energy.2019.05.138.
[74] Z. Wang, T. Hong, M.A. Piette, Building thermal load prediction through shallow machine learning and deep learning, Appl. Energy 263 (2020), https://doi.org/10.1016/j.apenergy.2020.114683 114683.
[75] A. Rahman, A.D. Smith, Predicting heating demand and sizing a stratified thermal storage tank using deep learning algorithms, Appl. Energy 228 (2018) 108–121, https://doi.org/10.1016/j.apenergy.2018.06.064.
[76] C. Zhang, J. Li, Y. Zhao, T. Li, Q. Chen, X. Zhang, A hybrid deep learning-based method for short-term building energy load prediction combined with an interpretation process, Energy Build. 225 (2020), https://doi.org/10.1016/j.enbuild.2020.110301 110301.
[77] L. Wen, K. Zhou, S. Yang, Load demand forecasting of residential buildings using a deep learning model, Electr. Power Syst. Res. 179 (2020), https://doi.org/10.1016/j.epsr.2019.106073 106073.
[78] R. Sendra-Arranz, A. Gutiérrez, A long short-term memory artificial neural network to predict daily HVAC consumption in buildings, Energy Build. 216 (2020), https://doi.org/10.1016/j.enbuild.2020.109952 109952.
[79] C. Fan, J. Wang, W. Gang, S. Li, Assessment of deep recurrent neural network-based strategies for short-term building energy predictions, Appl. Energy 236 (2019) 700–710, https://doi.org/10.1016/j.apenergy.2018.12.004.
[80] S. Ben Taieb, G. Bontempi, A.F. Atiya, A. Sorjamaa, A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition, Expert Syst. Appl. (2012), https://doi.org/10.1016/j.eswa.2012.01.039.
[81] E. Skomski, J.Y. Lee, W. Kim, V. Chandan, S. Katipamula, B. Hutchinson, Sequence-to-sequence neural networks for short-term electrical load forecasting in commercial office buildings, Energy Build. 226 (2020), https://doi.org/10.1016/j.enbuild.2020.110350.
[82] L. Sehovac, K. Grolinger, Deep Learning for Load Forecasting: Sequence to Sequence Recurrent Neural Networks with Attention, IEEE Access 8 (2020) 36411–36426, https://doi.org/10.1109/ACCESS.2020.2975738.
[83] D. Bahdanau, K. Cho, Y. Bengio, Neural Machine Translation by Jointly Learning to Align and Translate, (2014). http://arxiv.org/abs/1409.0473.
[84] M.-T. Luong, H. Pham, C.D. Manning, Effective Approaches to Attention-based Neural Machine Translation, n.d.
[85] C. Gallicchio, A. Micheli, L. Pedrelli, Deep reservoir computing: A critical experimental analysis, Neurocomputing. 268 (2017) 87–99, https://doi.org/10.1016/j.neucom.2016.12.089.
[86] C. Gallicchio, A. Micheli, L. Pedrelli, Design of deep echo state networks, Neural Networks. 108 (2018) 33–47, https://doi.org/10.1016/j.neunet.2018.08.002.
[87] G. Shi, D. Liu, Q. Wei, Energy consumption prediction of office buildings based on echo state networks, Neurocomputing. 216 (2016) 478–488, https://doi.org/10.1016/j.neucom.2016.08.004.
[88] M. Mansoor, F. Grimaccia, S. Leva, M. Mussetta, Comparison of echo state network and feed-forward neural networks in electrical load forecasting for demand response programs, Math. Comput. Simul (2020), https://doi.org/10.1016/j.matcom.2020.07.011.

[89] L. Wang, H. Hu, X.Y. Ai, H. Liu, Effective electricity energy consumption forecasting using echo state network improved by differential evolution algorithm, Energy. 153 (2018) 801–815, https://doi.org/10.1016/j.energy.2018.04.078.
[90] H. Hu, L. Wang, S.X. Lv, Forecasting energy consumption and wind power generation using deep echo state network, Renewable Energy 154 (2020) 598–613, https://doi.org/10.1016/j.renene.2020.03.042.
[91] H. Hu, L. Wang, L. Peng, Y.R. Zeng, Effective energy consumption forecasting using enhanced bagged echo state network, Energy. 193 (2020), https://doi.org/10.1016/j.energy.2019.116778 116778.
[92] G. Fud, Deep belief network based ensemble approach for cooling load forecasting of air-conditioning system, Energy. 148 (2018) 269–282, https://doi.org/10.1016/j.energy.2018.01.180.
[93] G. Hafeez, K.S. Alimgeer, I. Khan, Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid, Appl. Energy 269 (2020), https://doi.org/10.1016/j.apenergy.2020.114915 114915.
[94] J. Gamboa, Deep Learning for Time-Series Analysis, ArXiv. (2017).
[95] H.J. Sadaei, P.C. de Lima e Silva, F.G. Guimarães, M.H. Lee, Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series, Energy 175 (2019) 365–377, https://doi.org/10.1016/j.energy.2019.03.081.
[96] M. Cai, M. Pipattanasomporn, S. Rahman, Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques, Appl. Energy 236 (2019) 1078–1088, https://doi.org/10.1016/j.apenergy.2018.12.042.
[97] T.Y. Kim, S.B. Cho, Predicting residential energy consumption using CNN-LSTM neural networks, Energy. 182 (2019) 72–81, https://doi.org/10.1016/j.energy.2019.05.230.
[98] M. Sajjad, Z.A. Khan, A. Ullah, T. Hussain, W. Ullah, M.Y. Lee, S.W. Baik, A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting, IEEE Access 8 (2020) 143759–143768, https://doi.org/10.1109/ACCESS.2020.3009537.
[99] J. Donahue, L.A. Hendricks, M. Rohrbach, S. Venugopalan, S. Guadarrama, K. Saenko, T. Darrell, Long-Term Recurrent Convolutional Networks for Visual Recognition and Description, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2017) 677–691, https://doi.org/10.1109/TPAMI.2016.2599174.
[100] X. Shi, Z. Chen, H. Wang, D.Y. Yeung, W.K. Wong, W.C. Woo, Convolutional LSTM network: A machine learning approach for precipitation nowcasting, Adv. Neural Inform. Process. Syst. 2015 (2015) 802–810.
[101] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015, 10.1109/CVPR.2015.7298594.
[102] L.T. Le, H. Nguyen, J. Dou, J. Zhou, A comparative study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for smart city planning, Appl. Sci. (Switzerland). (2019), https://doi.org/10.3390/app9132630.
[103] S. Bouktif, A. Fiaz, A. Ouni, M.A. Serhani, Multi-sequence LSTM-RNN deep learning and metaheuristics for electric load forecasting, Energies (2020), https://doi.org/10.3390/en13020391.
[104] H. Su, E. Zio, J. Zhang, M. Xu, X. Li, Z. Zhang, A hybrid hourly natural gas demand forecasting method based on the integration of wavelet transform and enhanced Deep-RNN model, Energy. (2019), https://doi.org/10.1016/j.energy.2019.04.167.
[105] M. Ilbeigi, M. Ghomeishi, A. Dehghanbanadaki, Prediction and optimization of energy consumption in an office building using artificial neural network and a genetic algorithm, Sustain. Cities Soc. (2020), https://doi.org/10.1016/j.scs.2020.102325.
[106] 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, Renew. Sustain. Energy Rev. (2020), https://doi.org/10.1016/j.rser.2020.109980.
[107] G. Zhou, H. Moayedi, M. Bahiraei, Z. Lyu, Employing artificial bee colony and particle swarm techniques for optimizing a neural network in prediction of heating and cooling loads of residential buildings, J. Cleaner Prod. (2020), https://doi.org/10.1016/j.jclepro.2020.120082.
[108] J.R.S. Iruela, L.G.B. Ruiz, M.C. Pegalajar, M.I. Capel, A parallel solution with GPU technology to predict energy consumption in spatially distributed buildings using evolutionary optimization and artificial neural networks, Energy Convers. Manage. (2020), https://doi.org/10.1016/j.enconman.2020.112535.
[109] G. Zhou, H. Moayedi, L.K. Foong, Teaching–learning-based metaheuristic scheme for modifying neural computing in appraising energy performance of building, Eng. Comput. (2020), https://doi.org/10.1007/s00366-020-00981-5.
[110] A.S. Shah, H. Nasir, M. Fayaz, A. Lajis, I. Ullah, A. Shah, Dynamic user preference parameters selection and energy consumption optimization for smart homes using deep extreme learning machine and bat algorithm, IEEE Access (2020), https://doi.org/10.1109/access.2020.3037081.
[111] J. Wang, W. Yang, P. Du, Y. Li, Research and application of a hybrid forecasting framework based on multi-objective optimization for electrical power system, Energy. (2018), https://doi.org/10.1016/j.energy.2018.01.112.
[112] L. Wang, S.X. Lv, Y.R. Zeng, Effective sparse adaboost method with ESN and FOA for industrial electricity consumption forecasting in China, Energy. (2018), https://doi.org/10.1016/j.energy.2018.04.175.
[114] A. Naik, S.C. Satapathy, A. Abraham, Modified Social Group Optimization—a meta-heuristic algorithm to solve short-term hydrothermal scheduling, Appl. Soft Comput. J. 95 (2020), https://doi.org/10.1016/j.asoc.2020.106524.
[115] P. Lu, L. Ye, Y. Zhao, B. Dai, M. Pei, Y. Tang, Review of meta-heuristic algorithms for wind power prediction: Methodologies, applications and challenges, Appl. Energy 301 (2021), https://doi.org/10.1016/j.apenergy.2021.117446.
[116] R.P. Parouha, P. Verma, State-of-the-art reviews of meta-heuristic algorithms with their novel proposal for unconstrained optimization and applications, Arch. Comput. Methods Eng. 28 (2021) 4049–4115, https://doi.org/10.1007/s11831-021-09532-7.
[117] M.N. Fekri, H. Patel, K. Grolinger, V. Sharma, Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network, Appl. Energy 282 (2021), https://doi.org/10.1016/j.apenergy.2020.116177.
[118] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. (2010), https://doi.org/10.1109/TKDE.2009.191.
[119] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, A comprehensive survey on transfer learning, Proc. IEEE (2021), https://doi.org/10.1109/JPROC.2020.3004555.
[120] X. Fang, G. Gong, G. Li, L. Chun, W. Li, P. Peng, A hybrid deep transfer learning strategy for short term cross-building energy prediction, Energy 215 (2021), https://doi.org/10.1016/j.energy.2020.119208 119208.
[121] Z. Jiang, Y.M. Lee, Deep Transfer Learning for Thermal Dynamics Modeling in Smart Buildings, in: Proceedings - 2019 IEEE International Conference on Big Data, Big Data 2019, 2019, pp. 2033–2037, https://doi.org/10.1109/BigData47090.2019.9006306.
[122] M. Ribeiro, K. Grolinger, H.F. ElYamany, W.A. Higashino, M.A.M. Capretz, Transfer learning with seasonal and trend adjustment for cross-building energy forecasting, Energy Build. 165 (2018) 352–363, https://doi.org/10.1016/j.enbuild.2018.01.034.
[123] C. Fan, Y. Sun, F. Xiao, J. Ma, D. Lee, J. Wang, Y.C. Tseng, Statistical investigations of transfer learning-based methodology for short-term building energy predictions, Appl. Energy 262 (2020), https://doi.org/10.1016/j.apenergy.2020.114499 114499.
[124] Y. Gao, Y. Ruan, C. Fang, S. Yin, Deep learning and transfer learning models of energy consumption forecasting for a building with poor information data, Energy Build. 223 (2020), https://doi.org/10.1016/j.enbuild.2020.110156 110156.
[125] W. Dai, Q. Yang, G.R. Xue, Y. Yu, Boosting for transfer learning, ACM Internat. Conf. Proc. Series (2007), https://doi.org/10.1145/1273496.1273521.
[126] K.M. Borgwardt, A. Gretton, M.J. Rasch, H.P. Kriegel, B. Schölkopf, A.J. Smola, Integrating structured biological data by Kernel Maximum Mean Discrepancy, Bioinformatics (2006), https://doi.org/10.1093/bioinformatics/btl242.
[127] J. Davis, P. Domingos, Deep transfer via second-order Markov logic, in: Proceedings of the 26th International Conference On Machine Learning 2009, 2009, pp. 217–224.
[128] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Adv. Neural Inform. Process. Syst. (2014), https://doi.org/10.3156/jsoft.29.5_177_2.
[129] Y. Pang, X. Zhou, D. Xu, Z. Tan, M. Zhang, N. Guo, Y. Tian, Generative adversarial learning based commercial building electricity time series prediction, in: Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI. 2019-November, 2019, pp. 1800–1804, 10.1109/ICTAI.2019.00271.
[130] Z. Wang, T. Hong, Generating realistic building electrical load profiles through the Generative Adversarial Network (GAN), Energy Build. 224 (2020), https://doi.org/10.1016/j.enbuild.2020.110299 110299.
[131] M.N. Fekri, A.M. Ghosh, K. Grolinger, Generating energy data for machine learning with recurrent generative adversarial networks, Energies. 13 (2019), https://doi.org/10.3390/en13010130.
[132] Y.H. Kwon, M.G. Park, Predicting future frames using retrospective cycle gan, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2019-June, 2019, pp. 1811–1820, https://doi.org/10.1109/CVPR.2019.00191.

Further reading

[113] I. Fister, X.S. Yang, J. Brest, D. Fister, A brief review of nature-inspired algorithms for optimization, Elektrotehniski Vestnik/Electrotechnical Review (2013).
[133] A. Koochali, P. Schichtel, A. Dengel, S. Ahmed, Probabilistic forecasting of sensory data with generative adversarial networks - ForGAN, IEEE Access (2019), https://doi.org/10.1109/ACCESS.2019.2915544.
[134] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, A. Alahi, Social GAN: Socially acceptable trajectories with generative adversarial networks, ArXiv (2018).