Neurocomputing 216 (2016) 478–488

Contents lists available at ScienceDirect

Neurocomputing

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

Energy consumption prediction of office buildings based on echo state networks

Guang Shi [a], Derong Liu [b], Qinglai Wei [a,*]

[a] The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

ARTICLE INFO

Article history:
Received 3 December 2015
Received in revised form
1 June 2016
Accepted 1 August 2016
Communicated by H. Zhang
Available online 8 August 2016

Keywords:
Energy consumption
Time-series prediction
Office buildings
Echo state networks
Reservoir topologies

ABSTRACT

In this paper, energy consumption of an office building is predicted based on echo state networks (ESNs). Energy consumption of the office building is divided into consumptions from sockets, lights and air-conditioners, which are measured in each room of the office building by three ammeters installed inside, respectively. On the other hand, an office building generally consists of several types of rooms, i.e., office rooms, computer rooms, storage rooms, meeting rooms, etc., the energy consumption of which varies in accordance with different working routines in each type of rooms. In this paper, several novel reservoir topologies of ESNs are developed, the performance of ESNs with different reservoir topologies in predicting the energy consumption of rooms in the office building is compared, and the energy consumption of all the rooms in the office building is predicted with the developed topologies. Moreover, parameter sensitivity of ESNs with different reservoir topologies is analyzed. A case study shows that the developed simplified reservoir topologies are sufficient to achieve outstanding performance of ESNs in the prediction of building energy consumption.

1. Introduction

The increasing population, growing industrial production, fast economic development and rapid social progress in recent years have brought about huge demands for energy supplies and constantly rising energy consumption across the world, thereby resulting in a large number of environmental problems, including air pollution, water contamination, greenhouse effect, etc. [1,2]. In terms of energy consumption, buildings have become a focus of energy policy and decision making, since energy consumption from buildings accounts for a significant proportion throughout the world [3]. The proportion of building energy consumption in total energy consumption is approaching 40% in Europe [4], while it was 28% in China in 2011 and is expected to reach 35% by 2020 [5]. In order for decision makers to develop and implement policies to effectively reduce energy consumption from buildings, it is of great importance to establish accurate prediction of building energy consumption, with a view to alleviating environmental pollution to a certain extent and achieving sustainable development of human society. However, prediction of energy consumption is a great challenge due to factors including weather conditions, building structure, geographic location, settled population, seasonal changes, etc. [6,7].

During the past decades, a variety of techniques were applied to the prediction of building energy consumption, e.g., engineering methods [8], data mining techniques [9], neural networks (NNs) [10], clustering analysis [11], support vector machine (SVM) [12], fuzzy logic [13], etc. Among existing methods, NNs have been widely studied and applied to various fields, including system modeling [14], optimal control [15], fault diagnosis [16], and adaptive dynamic programming [17]. NNs are commonly divided into feedforward neural networks (FNNs) and recurrent neural networks (RNNs). Specifically, RNNs, which were widely used in nonlinear time-series prediction [18–21], have demonstrated their application to the prediction of building energy consumption [22]. However, traditional RNNs suffer from a high computational complexity in their training, which may lead to slow training, complex performance surfaces, and possible instability [23].

In recent years, echo state networks (ESNs), as a new type of RNNs proposed by Jaeger et al. [24,25], attracted great attention among researchers [26–31]. By using Markovian architectural bias of untrained RNNs to reflect historical inputs, ESNs utilize the dynamics created by a huge randomly created layer of recurrent units called reservoir, and only the connections between the reservoir and the output layer are modified in the learning process. In this way, the high computational complexity of traditional RNNs is significantly reduced, and the training efficiency is remarkably

* Corresponding author.
E-mail addresses: shiguang2012@ia.ac.cn (G. Shi), derong@ustb.edu.cn (D. Liu), qinglai.wei@ia.ac.cn (Q. Wei).

http://dx.doi.org/10.1016/j.neucom.2016.08.004
0925-2312/

enhanced. The past decades have witnessed extensive studies on ESNs both in theory and practice.

Theoretical studies on ESNs focused on reservoir optimization mainly carried out from structure improvement and parameter selection. In [32], multiple network topologies were used for the generation of reservoir, including small-world network, scale-free network, etc. Deng and Zhang [33] proposed a reservoir topology with small-world and scale-free properties. Song and Feng [34] investigated a cortex-like network generation method to construct the reservoir and therefore discovered an improved design strategy for the reservoir. Rodan and Tino [35] developed three simple reservoir topologies called delay line reservoir (DLR), DLR with feedback connections (DLRB) and simple cycle reservoir (SCR), respectively, which achieved good performance in time-series processing. In [36], a stochastic gradient descent method was developed to optimize global learning parameters including input and output feedback scalings, leaking rate and spectral radius of ESNs. Steil [37] optimized the reservoir by adopting a biologically motivated learning rule based on neural intrinsic plasticity (IP). In [38,39], the learning rule of IP was applied to tuning the probability density of all the neurons' outputs towards an exponential distribution and Gaussian distribution, respectively, which realized maximization of information. On the other hand, ESNs have achieved wide practical applications in various fields, including chaotic time-series prediction [40–42], dynamic pattern extraction [43], speech recognition [44], noise modeling [25] and complex signal filtering [45]. Specifically, ESNs have demonstrated outstanding performance in time-series prediction with real-life measurements [46–49].

Inspired by [35], several simplified reservoir topologies of ESNs are developed in this paper, and then ESNs with these different reservoir topologies are applied to the prediction of building energy consumption, which fully utilizes the remarkable performance of ESNs in chaotic time-series prediction. The main contributions of this paper are summarized as follows:

(1) A total of six simplified reservoir topologies of ESNs including two novel ones are developed and described in detail.

(2) The energy consumption in an office building is predicted by ESNs with different reservoir topologies, and their prediction performance is compared.

(3) The parameters of different topologies are summarized, and a sensitivity analysis is conducted to show the influence of these parameters on the prediction performance of the ESNs.

The rest of the paper is organized as follows. Basic knowledge on ESNs is given in Section 2. Several different reservoir topologies of ESNs are developed in Section 3, and related parameters of the topologies are summarized. A detailed case study with a parameter sensitivity analysis is given in Section 4 to show the performance of the developed topologies in the prediction of building energy consumption. Finally, the conclusion is drawn and future work is presented in Section 5.

2. Preliminaries of ESNs

As shown in Fig. 1, an ESN is a discrete-time RNN composed of an input layer, a reservoir and an output layer. The reservoir contains a large number of interconnected dynamic units, which are called reservoir units in this paper. The output layer is a memoryless linear readout trained to generate the output.

It is assumed that an ESN has K input units, N reservoir units and L output units. Activations of the input, reservoir and output units at time step t are denoted by s_t, x_t and o_t, respectively, where $s_t \in R^K$, $x_t \in R^N$ and $o_t \in R^L$. Connections between input units and reservoir units are collected in an $N \times K$ weight matrix W^{in}, connections between reservoir units are collected in an $N \times N$ weight matrix W, while connections between reservoir units and output units are collected in an $N \times L$ weight matrix W^{out}. Therefore, the activation of reservoir units is updated as

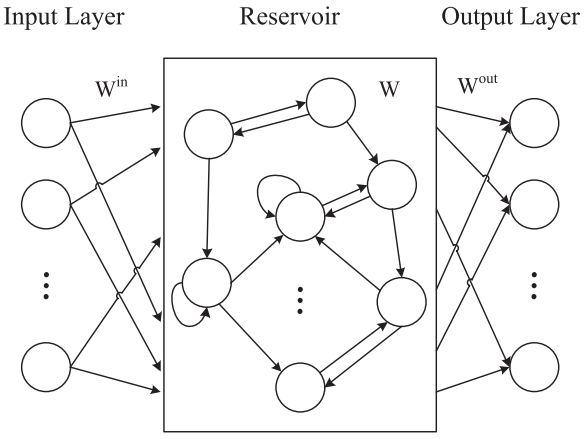

Fig. 1. Basic structure of an ESN.

$$x_t = f(W^{\text{in}} \cdot s_t + W \cdot x_{t-1}), \tag{1}$$

where $f = \tanh$ is the activation function of the reservoir. The linear output is calculated as

$$o_t = (W^{\text{out}})^T \cdot x_t. \tag{2}$$

Elements of W^{in} and W are randomly initialized from a continuous probability distribution before training. In order to guarantee the echo state property (ESP) of reservoir units, the reservoir weight matrix W is usually scaled by $\eta W/\rho \to W$, where $0 < \eta < 1$ is a scaling parameter and ρ is the spectral radius of W.

Remark 1. As given in [50,51], setting the spectral radius of W less than 1 is only a necessary condition to guarantee the ESP of reservoir units, however, the necessary condition is often sufficient in most practical cases owing to the contractive dynamics of the reservoir with the nonlinear activation function $f = \tanh$.

During the training of the ESN, W^{in} and W are fixed, and only W^{out} is tuned. Therefore, given M training samples and based on (1) and (2), we can obtain the following equation

$$\mathcal{H} W^{\text{out}} = O, \tag{3}$$

where the matrix $\mathcal{H}$ called reservoir output matrix is constructed as

$$\mathcal{H}\left(w_1^{\text{in}}, \ldots, w_N^{\text{in}}, w_1, \ldots, w_N, s_1, \ldots, s_M, x_0, \ldots, x_{M-1}\right) = \begin{bmatrix} f\left(w_1^{\text{in}} \cdot s_1 + w_1 \cdot x_0\right) & \cdots & f\left(w_N^{\text{in}} \cdot s_1 + w_N \cdot x_0\right) \\ \vdots & \ddots & \vdots \\ f\left(w_1^{\text{in}} \cdot s_M + w_1 \cdot x_0\right) & \cdots & f\left(w_N^{\text{in}} \cdot s_M + w_N \cdot x_{M-1}\right) \end{bmatrix}_{M \times N}, \tag{4}$$

and

$$W^{\text{out}} = \begin{bmatrix} \left(w_1^{\text{out}}\right)^T \\ \vdots \\ \left(w_N^{\text{out}}\right)^T \end{bmatrix}_{N \times L}, \quad \text{and} \quad O = \begin{bmatrix} o_1^T \\ \vdots \\ o_M^T \end{bmatrix}_{M \times L}. \tag{5}$$

We can calculate the minimum norm least-square solution of the linear system (3) as the output weight matrix of the ESN,

$$\hat{W}^{\text{out}} = \mathcal{H}^{\dagger} O, \tag{6}$$

where $\mathcal{H}^{\dagger}$ is the Moore–Penrose generalized inverse of matrix $\mathcal{H}$ [52,53].

Finally, we adopt three statistical indexes, namely mean square error (MSE), root mean square error (RMSE) and coefficient of variation of the root mean square error (CV-RMSE), to evaluate the training effect of an ESN, and their definitions are given as follows

$$\begin{aligned}
&\mathrm{MSE} = \frac{1}{M}\sum_{t=1}^{M}\left(o_t - \hat{o}_t\right)^2,\\
&\mathrm{RMSE} = \sqrt{\mathrm{MSE}},\\
&\text{CV-RMSE} = \frac{\mathrm{RMSE}}{A}\times 100\%,\\
&A = \frac{1}{M}\sum_{1}^{M} o_t,
\end{aligned} \tag{7}$$

where $\hat{o}_t$ is the network input, o_t is the target output, and M is the number of training samples.

3. Reservoir topologies

In this section, several simplified reservoir topologies are developed including those in [35]. We aim to compare the newly developed topologies with existing ones with respect to performance in the prediction of building energy consumption.

3.1. Reservoir topologies

Apart from the traditional ESN reservoir described in Section 2, six simplified reservoir topologies as follows are developed, as shown in Fig. 2.

3.1.1. Non-connected reservoir (NCR)

Composed of reservoir units, none of which are connected with each other, i.e., all the elements in the reservoir weight matrix W are zero. Specifically, it is exactly the same as the well-known feedforward neural network (FNN).

3.1.2. Delay line reservoir (DLR)

With units arranged in a line, and only elements on the lower subdiagonal of the reservoir weight matrix W have non-zero values $W_{i+1,i} = r$ for $i = 1, \ldots, N-1$, where r is the weight of all the feedforward connections.

3.1.3. DLR with feedback connections (DLRB)

With the same structure as DLR, but each reservoir unit is also connected to the previous one. Elements on the lower and upper subdiagonals of the reservoir weight matrix W have non-zero values, i.e., $W_{i+1,i} = r$ and $W_{i,i+1} = b$, where b is the weight of all the feedback connections.

3.1.4. Simple cycle reservoir (SCR)

Consisting of units arranged in a cycle. Elements on the lower subdiagonal and at the upper-right corner of W are non-zero with $W_{i+1,i} = r$ and $W_{1,N} = r$.

3.1.5. Self-feedback DLR (SDLR)

Besides the same structure as DLR, each reservoir unit is connected to itself. In addition to non-zero elements on the lower subdiagonal of W, elements on the diagonal of W are non-zero as well, i.e., $W_{i,i} = d$ for $i = 1, \ldots, N$, where d is the weight of all the self-feedback connections.

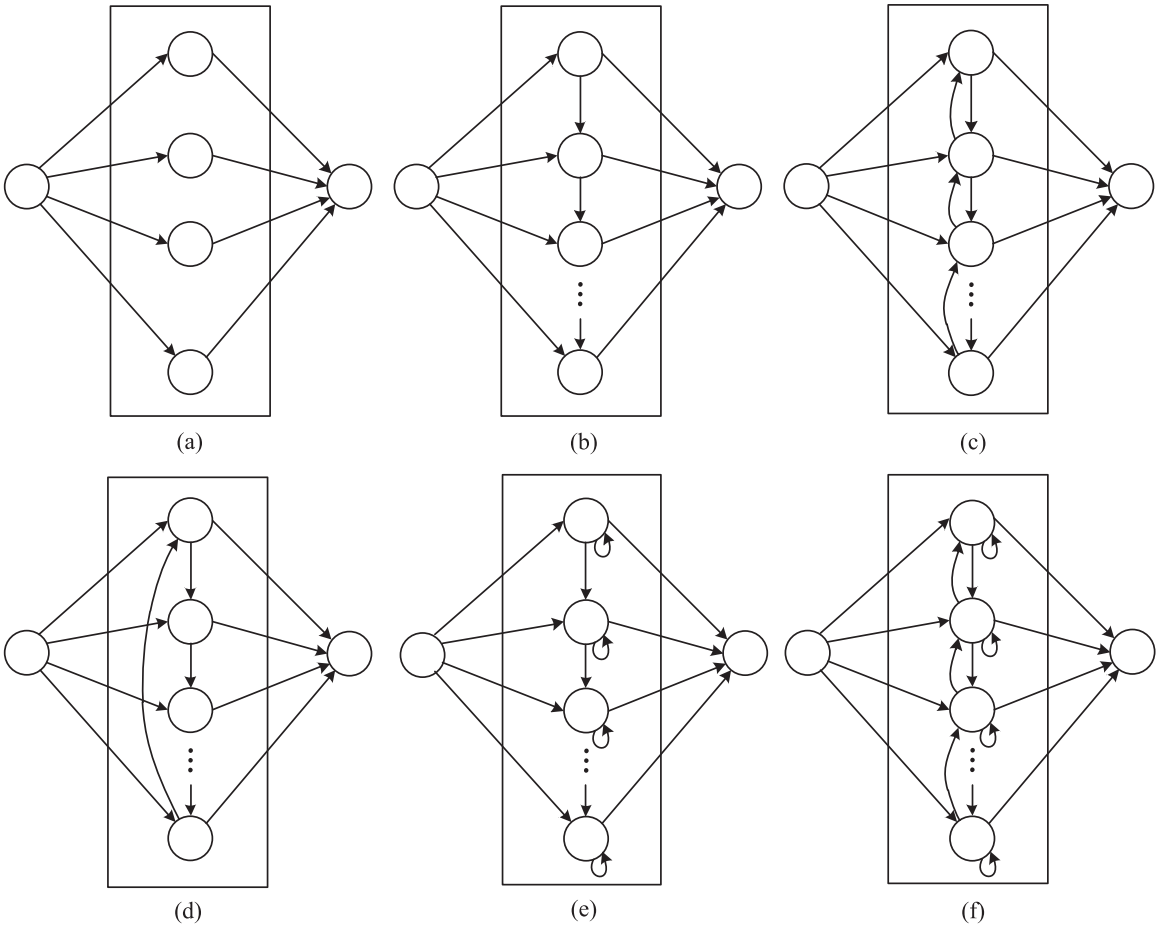

Fig. 2. Reservoir topologies. (a) NCR. (b) DLR. (c) DLRB. (d) SCR. (e) SDLR. (f) SDLRB.

3.1.6. Self-feedback DLRB (SDLRB)

Structured the same as DLRB with each reservoir unit connected to itself. Besides non-zero elements on the lower and upper subdiagonals of W, the diagonal of W has non-zero elements as $W_{i,i} = d$.

Remark 2. Among the six reservoir topologies above, NCR equals a traditional FNN, DLR, DLRB and SCR were given in [35], and SDLR and SDLRB are newly developed in this paper. In the next section, we aim to compare the performance of SDLR and SDLRB in energy consumption prediction with that of existing reservoir topologies.

3.2. Topology parameters

All the above topologies contain one or more parameters, which are summarized as follows.

3.2.1. Input connection v

In the topologies in Fig. 2, the input layer is fully connected to the reservoir. For traditional ESNs, the input weights are usually generated randomly from a uniform distribution over an interval $[-a, a]$. For the developed simple reservoirs (NCR, DLR, DLRB, SCR, SDLR and SDLRB), all the input connections have the same absolute weight value $v \in [0, 1]$, while the sign of each input weight is randomly determined.

3.2.2. Feedforward connection r

In the developed topologies (DLR, DLRB, SCR, SDLR and SDLRB), $r \in [0, 1]$ denotes the connection between each reservoir unit and its following one, namely the value of elements on the lower subdiagonal of the reservoir weight matrix W.

3.2.3. Feedback connection b

In DLRB and SDLRB, $b \in [0, 1]$ denotes the connection between each reservoir unit and its previous one, namely the value of elements on the upper subdiagonal of the reservoir weight matrix W.

3.2.4. Self-feedback connection d

In SDLR and SDLRB, $d \in [0, 1]$ denotes the connection between each reservoir unit and itself, namely the value of elements on the diagonal of the reservoir weight matrix W.

3.2.5. Reservoir size N

N, which denotes the number of units in the reservoir, has a significant influence on the training performance of ESNs.

Parameters of different topologies are summarized in Table 1. In the next section, ESNs with the above six reservoir topologies will be applied to the prediction of energy consumption in the rooms of an office building. The performance of ESNs with different reservoir topologies will be compared and a sensitivity analysis of the topology parameters will be given.

Table 1
Parameter summarization.

Topology	Parameter				
	v	r	b	d	N
NCR	*				*
DLR	*	*			*
DLRB	*	*	*		*
SCR	*	*			*
SDLR	*	*		*	*
SDLRB	*	*	*	*	*

4. Case study

In this section, a detailed case study is given to demonstrate the performance of ESNs with the developed reservoir topologies. The case study is based on an office building in one of our practical projects.

4.1. Background introduction

The building concerned in this paper is an office building located in a development zone in China and composed of a total of 14 floors with 6 rooms on each floor, excluding the first floor which is used as the lobby of the building. The building settles a certain number of companies engaged in information technology, each of which rent several rooms and use them as office rooms, computer rooms, storage rooms, meeting rooms, etc. A central air-conditioning system is adopted in the building, where air-conditioning in each room is controlled by several switches.

The rooms in the office building are summarized in Table 2, where we define office rooms as "⋆", computer rooms as "△", storage rooms as "⋄", and meeting rooms as "●". We can see that office rooms account for the majority of all the rooms in the building, so relevant results of an office room are mainly displayed next.

Three ammeters are installed in each room to measure real-time energy consumption from sockets, lights and air-conditioners at an interval of 1 h, while such three types of energy consumption can basically represent all the energy consumptions in a room. Given original data of energy consumption measured for 4 years from January 2011 to December 2014, the data are preprocessed to eliminate absence, incompleteness, repetition and other defects, thus obtaining complete hourly data of energy consumption. Next, because nobody works in the office building on weekends in our study, and therefore almost no energy consumption is generated then, the data from weekends are deleted. Furthermore, the remaining data on working days are randomly divided into training data and testing data with an approximate proportion of 7:3.

Specifically, we separate the prediction of energy consumption to different months, namely we focus on the data in the same month to carry out prediction. In the following, relevant results of energy consumption in July for 4 years are presented as an example. The average outside air temperature and building occupancy on the working days of July in each year are shown in Fig. 3 (a) and (b), respectively. It can be seen that both factors generally present stable fluctuations with few changes in different years with the temperature varying within the range of 23–26 °C and the occupancy 78–91%. Therefore, the impacts of these factors on the prediction could be ignored by using the method in this paper.

Table 2
Rooms in the office building.

Floor	Room					
	1	2	3	4	5	6
2	⋆	⋆	⋆	△	⋆	⋆
3	⋆	⋆	⋄	⋆	●	△
4	⋆	⋆	⋆	⋄	⋆	●
5	⋆	⋆	△	⋄	⋆	⋆
6	⋄	⋆	⋆	△	⋄	⋆
7	⋆	⋆	●	⋆	⋆	●
8	⋆	⋆	△	⋆	●	⋆
9	⋆	⋆	⋆	⋄	⋆	●
10	⋆	⋆	⋄	⋆	△	⋆
11	⋆	⋆	△	⋆	●	●
12	⋆	⋆	⋆	△	⋆	⋆
13	⋆	⋆	⋄	⋆	●	⋆
14	⋆	⋆	△	●	⋆	⋄

⋆: office room △: computer room ⋄: storage room ●: meeting room.

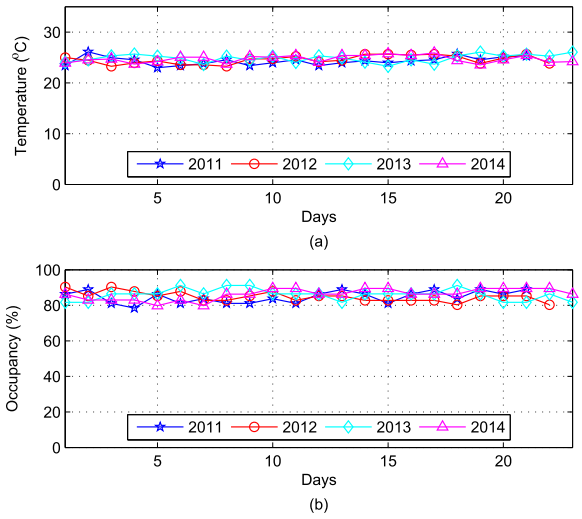

Fig. 3. (a) Average outside air temperature on the working days of July in 2011–2014. (b) Building occupancy on the working days of July in 2011–2014.

4.2. Office room

The original energy consumption of an office room in the building in 5 working days is shown as an example in Fig. 4, where the energy consumptions from sockets, lights and air-conditioners are shown by three curves, respectively. As indicated by Fig. 4, all three types of energy consumption present a periodic property with a typical "double-peak" characteristic; that is to say, all the three curves reach their peak values in mid-morning around 11:00 and mid-afternoon around 16:00 on a working day, and achieve a low point at noon since part of staff who often punctually go out for lunch then may switch off some electrical appliances using sockets, turn off some lights or adjust the temperature set for the air-conditioners, while some others who have their lunch inside the room may still consume some energy. Before and after work, however, due to no special requirements on energy consumption, all the appliances consuming energy are turned off when nobody stays in the room, so the energy consumption in non-working hours is close to zero. The means and variances of the entire time series are shown in Fig. 5.

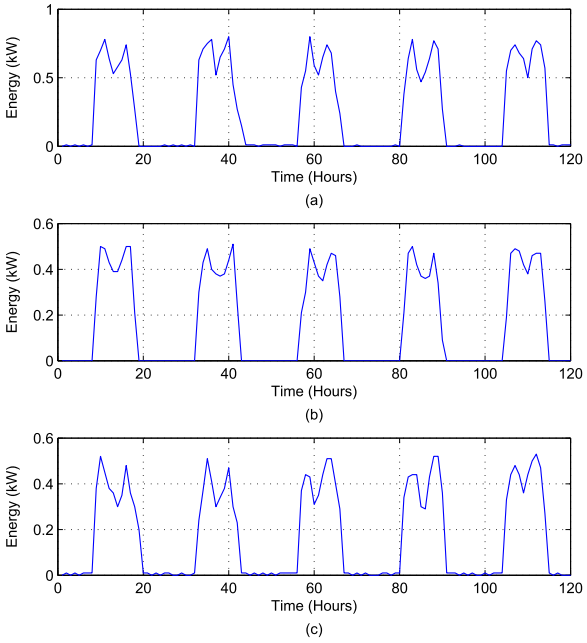

Fig. 4. Original energy consumption of the office room in 5 working days. (a) Energy consumption from sockets. (b) Energy consumption from lights. (c) Energy consumption from air-conditioners.

ESNs with different reservoir topologies are established and then trained with the training data. With a fixed reservoir size N each time, network parameters, including input connection $v \in [0, 1]$, feedforward connection $r \in [0, 1]$, feedback connection $b \in [0, 1]$ and self-feedback connection $d \in [0, 1]$, are randomly initialized from a uniform distribution over corresponding ranges in each training, respectively. After 100 times of training, we calculate the network outputs as the prediction results by the ESNs and regard the values of testing data as the real energy consumption. Next, we calculate the MSE, RMSE and CV-RMSE between predicted and real energy consumptions as the prediction errors of the ESNs. The results of energy consumption prediction in the office room by ESNs with different reservoir topologies are shown in Fig. 6, where only the average real energy and energy predicted by the newly developed SDLR and SDLRB are illustrated to simplify the display. The variances of the predicted and real time series are shown in Fig. 7. In case of negative prediction results, the results are adjusted to zero merely to satisfy practical conditions, but this adjustment is not considered in the calculation of prediction errors. Particularly, the reservoir topology of NCR, i.e., FNN, fails to achieve accurate prediction by using the training approach in this paper, so related results of NCR are omitted. In addition, the prediction errors of ESNs with different reservoir topologies are shown in Fig. 8.

As shown in Figs. 6–8, all the ESNs with different reservoir topologies achieve comparable performance in the energy consumption prediction of the office room with slight differences, while the training errors are high when the reservoir size N is small and low when N is large, but become stable when N is sufficiently large. The minimum prediction errors of ESNs with different reservoir topologies with $N=50$ are summarized in Table 3, where it can be seen that the traditional ESN generally achieves the minimum errors in the prediction, followed by SDLR and SDLRB which perform better than DLR, DLRB and SCR, and generally the more complex the reservoir topology, the smaller the prediction error.

Further, the original data of three types of energy consumption are added up to predict the entire energy consumption of the office room. By using the same approach above, the prediction results by SDLR and SDLRB are shown in Fig. 9, and the minimum prediction errors of the two topologies are summarized in Table 4.

4.3. Computer room

Generally speaking, a computer room is equipped with hosts, servers, switches and other computer equipments. The original energy consumption of a computer room in the building in 5 working days is shown in Fig. 10, where the energy consumptions from sockets, lights and air-conditioners are shown by three curves, respectively. The most remarkable difference between the computer room and the office room above lies in the energy consumption from sockets, which basically remains unchanged during a working day, since all the computer equipments consuming energy from sockets require 24-h stable running. However, for energy consumptions from lights and air-conditioners, both of them present the same "double-peak" characteristic as those in the office room considering similar working routines of staff in the computer room. Moreover, it is noteworthy that energy consumption from air-conditioners in the computer room remains at a constant non-zero value at night given the requirement on temperature from some computer equipments.

Similarly, we establish ESNs with different reservoir topologies and train them with the training data of the computer room.

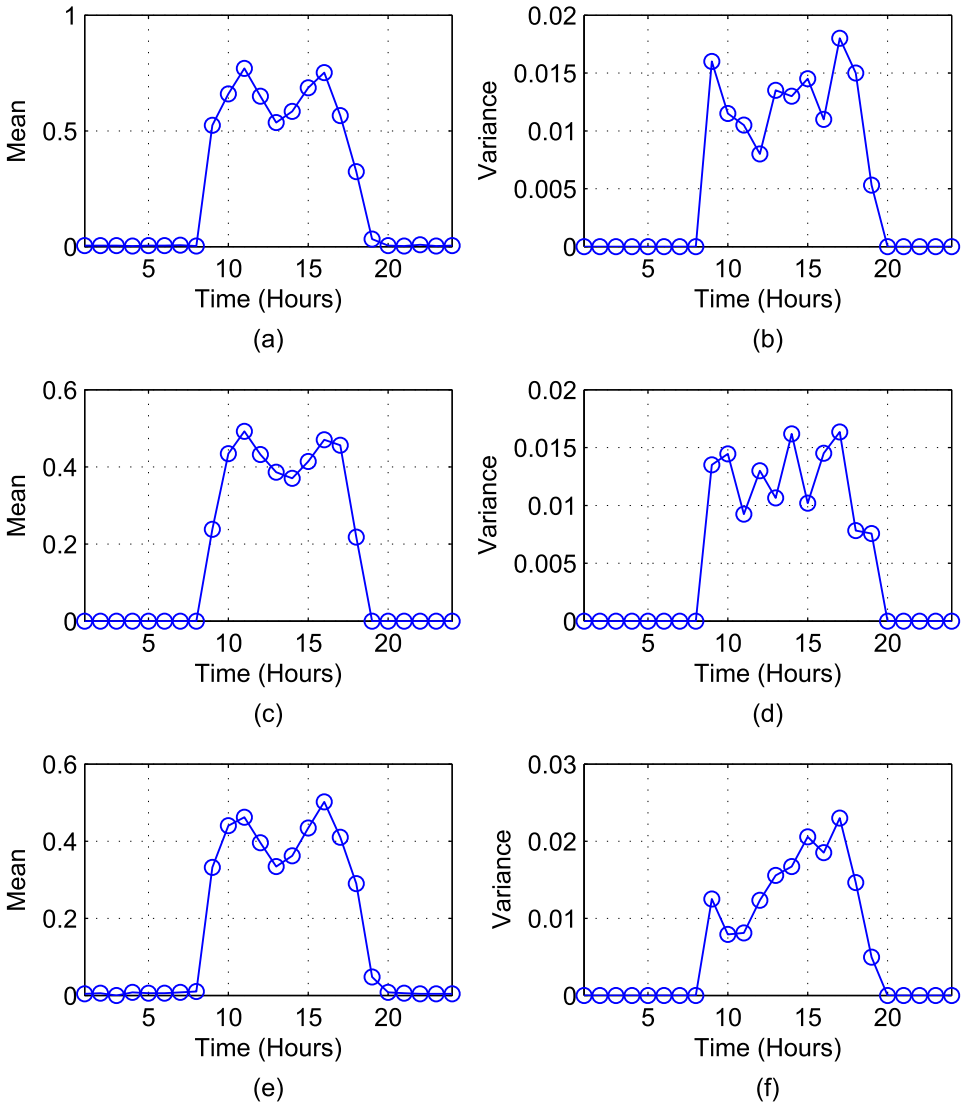

Fig. 5. Means and variances of energy consumption of the office room in July in 2011–2014. (a) Mean of energy consumption from sockets. (b) Variance of energy consumption from sockets. (c) Mean of energy consumption from lights. (d) Variance of energy consumption from lights. (e) Mean of energy consumption from air-conditioners. (f) Variance of energy consumption from air-conditioners.

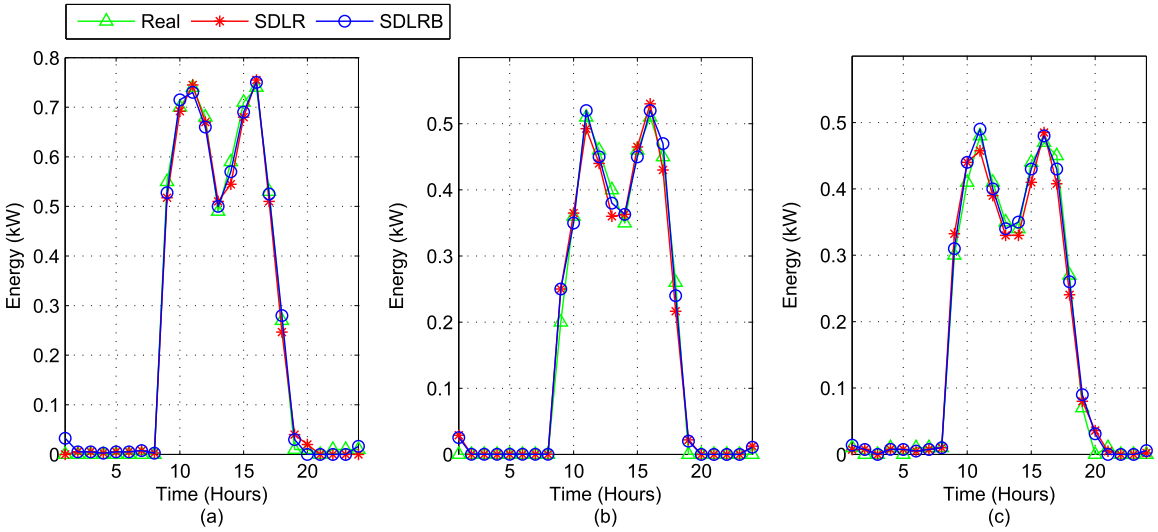

Fig. 6. Energy consumption prediction of the office room. (a) Prediction of energy consumption from sockets. (b) Prediction of energy consumption from lights. (c) Prediction of energy consumption from air-conditioners.

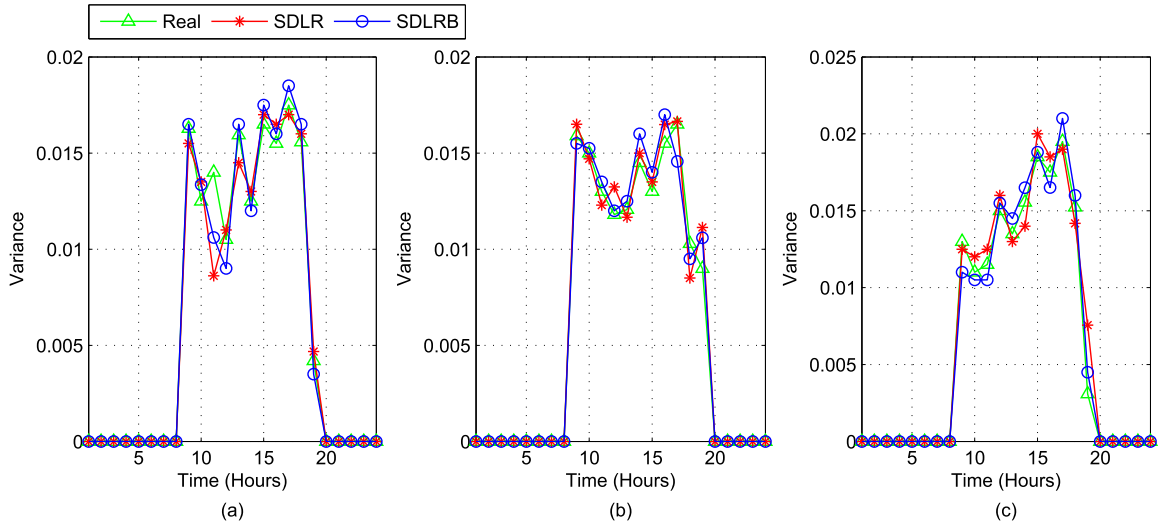

Fig. 7. Variances of predicted and real energy consumption of the office room. (a) Energy consumption from sockets. (b) Energy consumption from lights. (c) Energy consumption from air-conditioners.

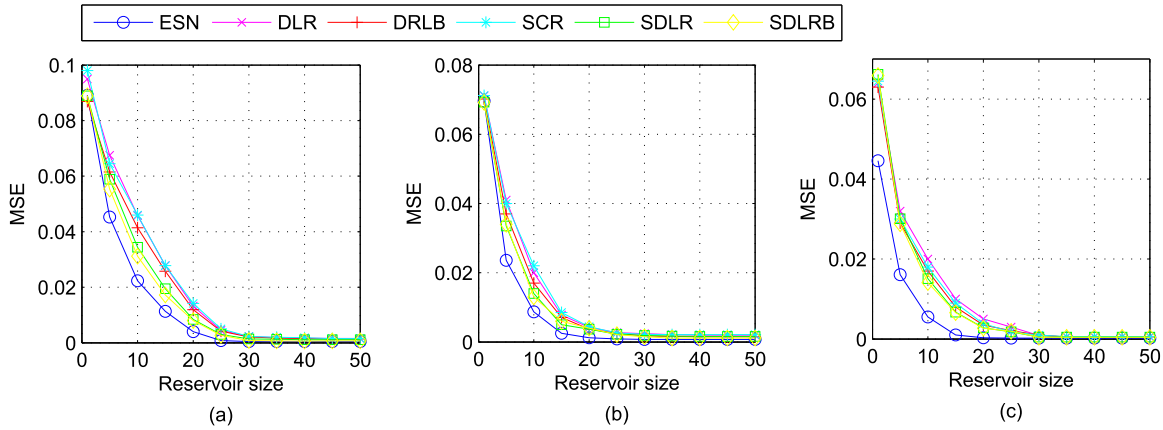

Fig. 8. Energy consumption prediction errors of the office room. (a) Prediction errors of energy consumption from sockets. (b) Prediction errors of energy consumption from lights. (c) Prediction errors of energy consumption from air-conditioners.

Table 3
Minimum errors of ESN, DLR, DLRB, SCR, SDLR and SDLRB for separate prediction of the office room with reservoir size $N=50$.

Consumption Type	Error	ESN	DLR	DLRB	SCR	SDLR	SDLRB
Sockets	MSE ($\times 10^{-4}$)	1.02	3.61	3.26	3.41	2.34	1.79
	RMSE ($\times 10^{-2}$)	1.01	1.90	1.81	1.85	1.53	1.34
	CV-RMSE (%)	4.00	7.54	7.16	7.33	6.07	5.31
Lights	MSE ($\times 10^{-4}$)	0.56	1.72	1.64	1.79	1.13	0.84
	RMSE ($\times 10^{-2}$)	0.75	1.31	1.28	1.34	1.06	0.92
	CV-RMSE (%)	4.52	7.94	7.77	8.10	6.43	5.55
Air-conditioners	MSE ($\times 10^{-4}$)	0.37	1.28	1.22	1.25	1.03	0.74
	RMSE ($\times 10^{-2}$)	0.61	1.13	1.10	1.12	1.01	0.86
	CV-RMSE (%)	3.62	6.70	6.54	6.62	6.00	5.09

Network parameters v, r, b and d are randomized in each training with a fixed N. After 100 times of training, network outputs are calculated as energy consumption predicted by the ESNs, and the values of testing data are regarded as real energy consumption. The results of energy consumption prediction in the computer room by ESNs with different reservoir topologies are shown in Fig. 11, which also only includes the real energy and energy predicted by SDLR and SDLRB. In view of similar effects with the office room, the curves and data of prediction errors are omitted here.

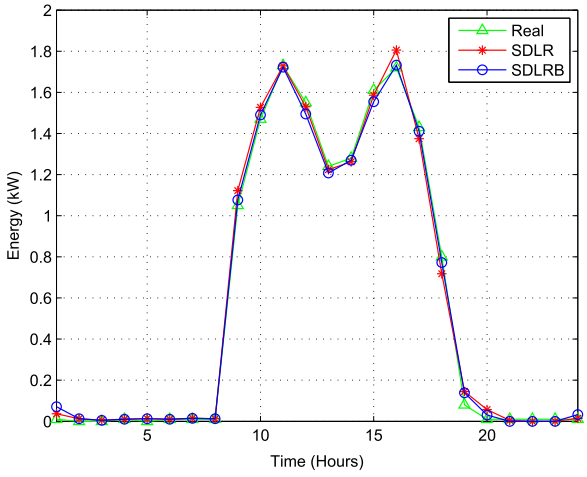

Fig. 9. Prediction of entire energy consumption in the office room.

4.4. Storage room

A storage room is generally stored with items requiring a constant temperature for storage. The original energy consumption of a storage room in the building in 5 working days is shown

Table 4
Minimum errors of SDLR and SDLRB for entire prediction of the office room with reservoir size $N=50$.

Error	SDLR	SDLRB
MSE ($\times 10^{-4}$)	11.55	7.69
RMSE ($\times 10^{-2}$)	3.40	2.77
CV-RMSE (%)	5.80	4.73

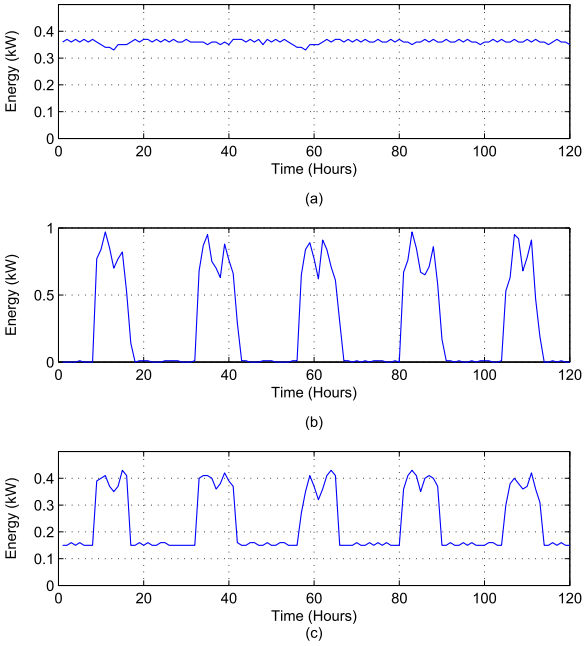

Fig. 10. Original energy consumption of the computer room in 5 working days. (a) Energy consumption from sockets. (b) Energy consumption from lights. (c) Energy consumption from air-conditioners.

in Fig. 12, where the energy consumptions from sockets, lights and air-conditioners are shown by three curves, respectively. It can be seen that all the three curves display entirely different characteristics, none of which still takes on the "double-peak" characteristic, but the energy consumption from air-conditioners remains constant in view of special storage requirements of items stored inside, while the other two curves of energy consumption are close

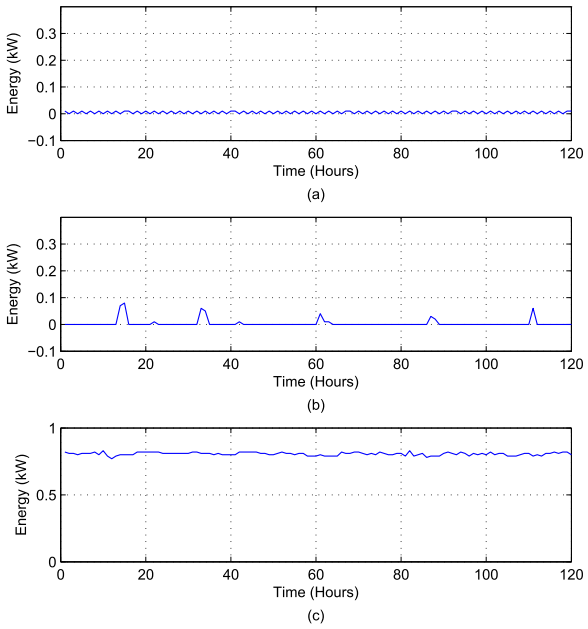

Fig. 12. Original energy consumption of the storage room in 5 working days. (a) Energy consumption from sockets. (b) Energy consumption from lights. (c) Energy consumption from air-conditioners.

to zero since nobody regularly works in the storage room, and therefore almost no energy consumption from sockets and lights is generated inside the room.

The energy consumption of the storage room is predicted with the same approach as above, and the prediction results only with real energy and energy predicted by SDLR and SDLRB are shown in Fig. 13.

4.5. Meeting room

The original energy consumption of a meeting room in the building in 5 working days is shown in Fig. 14, where the energy consumptions from sockets, lights and air-conditioners are shown by three curves, respectively. Since the meeting room is occasionally used without a fixed pattern, it can be seen that all the three curves of energy consumption are close to zero. Therefore, it is unnecessary to predict the energy consumption in the meeting room.

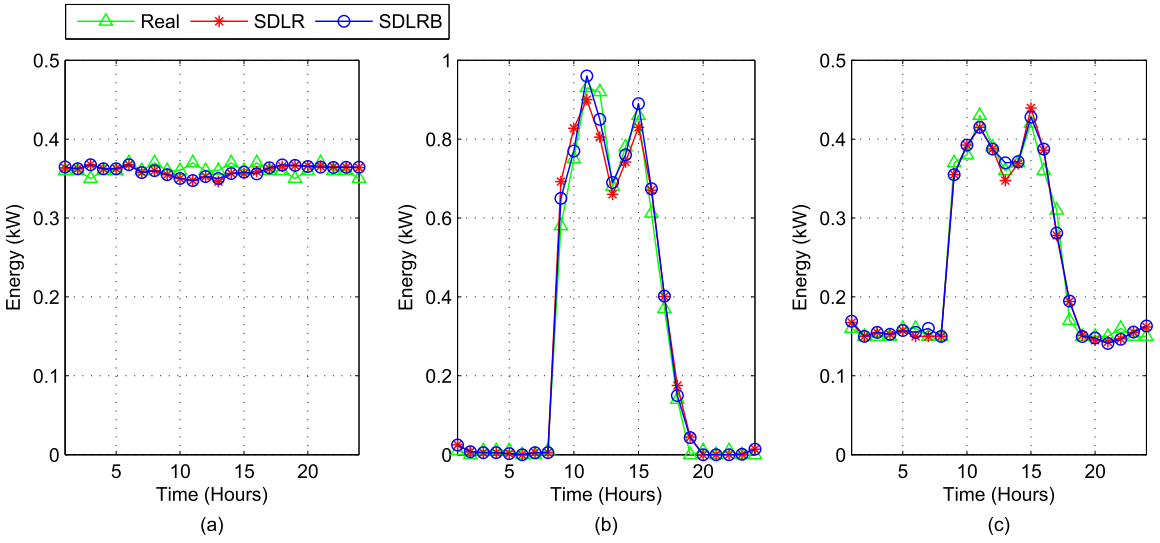

Fig. 11. Energy consumption prediction of the computer room. (a) Prediction of energy consumption from sockets. (b) Prediction of energy consumption from lights. (c) Prediction of energy consumption from air-conditioners.

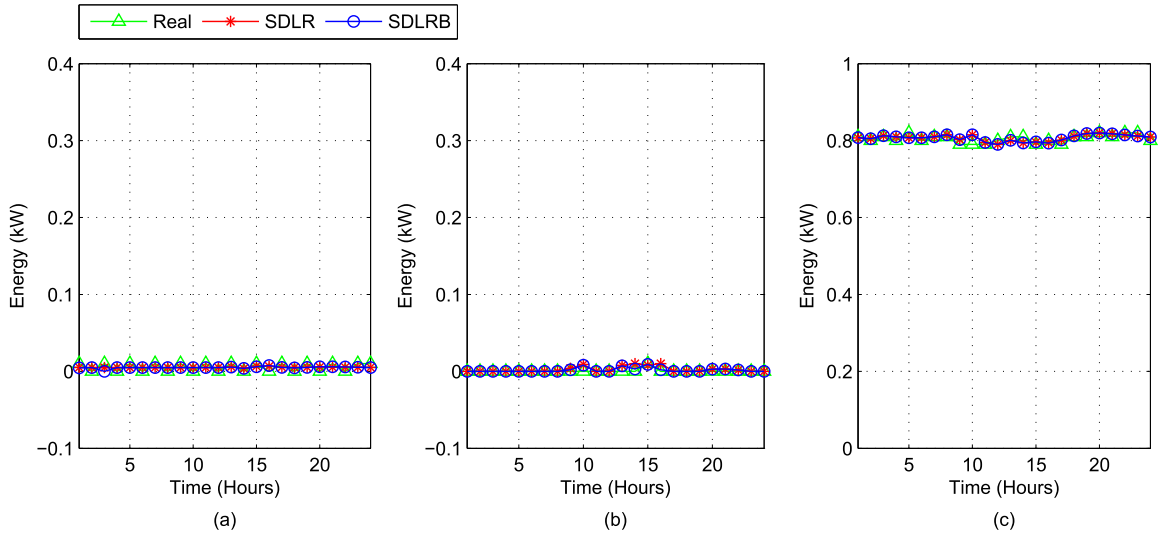

Fig. 13. Energy consumption prediction of the storage room. (a) Prediction of energy consumption from sockets. (b) Prediction of energy consumption from lights. (c) Prediction of energy consumption from air-conditioners.

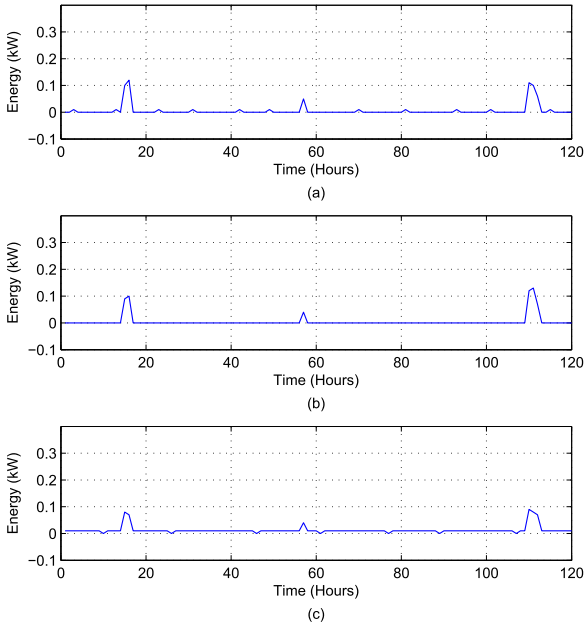

Fig. 14. Original energy consumption of the meeting room in 5 working days. (a) Energy consumption from sockets. (b) Energy consumption from lights. (c) Energy consumption from air-conditioners.

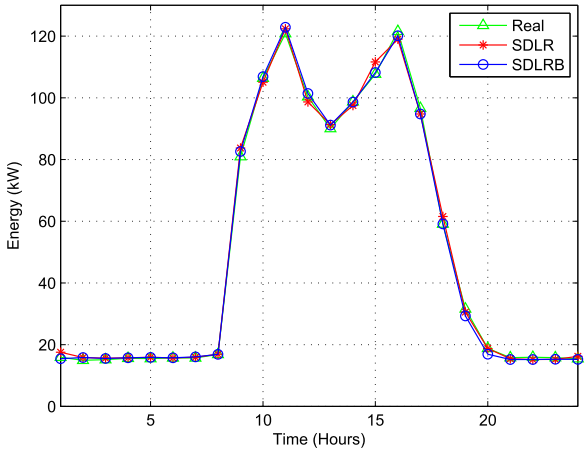

Fig. 15. Prediction of energy consumption in the whole office building.

4.6. Whole building

Finally, the energy consumption of all the rooms in the office building is added up and then predicted by the developed topologies SDLR and SDLRB. The results of prediction are shown in Fig. 15. The minimum CV-RMSEs of the two topologies are 4.97% and 3.72%, respectively, showing that the two topologies are sufficient to achieve high accuracy in the prediction of energy consumption of all the rooms in the whole building.

4.7. Sensitivity analysis

Furthermore, a sensitivity analysis on the ESNs with different topologies is conducted, the purpose of which is to evaluate the influence of topology parameters on the performance of ESNs in energy consumption prediction. In order to simplify the display, only the prediction of energy consumption from sockets in the aforementioned office room is given as a typical example. The reservoir size is set to $N=50$ for each ESN with a different reservoir topology. In each training, network parameters, including input connection $v \in [0, 1]$, feedforward connection $r \in [0, 1]$, feedback connection $b \in [0, 1]$ and self-feedback connection $d \in [0, 1]$, are incrementally initialized within corresponding ranges at an interval of 0.1 or fixed as appropriate. The results of sensitivity analysis are shown in Fig. 16. In respect of DLR, performance sensitivity of input connection v and feedforward connection r is shown in Fig. 16(a). For DLRB, Fig. 16(b) shows the performance sensitivity to changes in two parameters, namely the feedforward connection r and feedback connection b, while the input connection v is fixed. Sensitivity of the input connection v and feedforward connection r with fixed self-feedback connection d in SDLR is given in Fig. 16(c), and performance changes with respect to variations in the self-feedback connection d while the input connection v and feedforward connection r remain unchanged are shown in Table 5. For SDLRB, Fig. 16(d) illustrates the performance sensitivity of feedback connection b and self-feedback connection d with unchanged input connection v and feedforward connection r. In general, all the developed reservoir topologies present sufficient robustness with respect to small variations in the parameters.

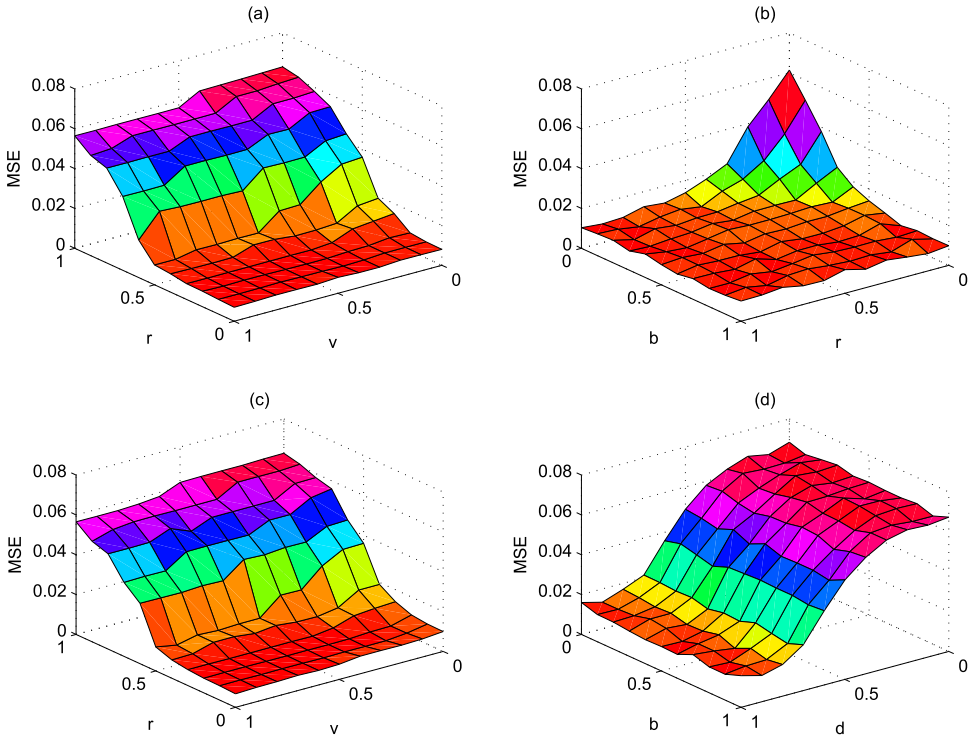

Fig. 16. Sensitivity analysis. (a) DLR. (b) DLRB. (c) SDLR. (d) SDLRB.

Table 5
Best input connection and feedforward connection for SDLR with different self-feedback connections in prediction of energy consumption from sockets in the office room.

v	r	d	MSE ($\times 10^{-4}$)
0.4	0.9	0.2	5.12
0.4	0.9	0.4	3.84
0.4	0.9	0.6	3.03
0.4	0.9	0.8	3.71
0.4	0.9	1	4.88

5. Conclusion and future work

In this paper, several simplified reservoir topologies of ESNs are developed and applied to the prediction of energy consumption in an office building, where rooms are classified into several types, including office rooms, computer rooms, storage rooms, meeting rooms, etc. The performance of ESNs with different reservoir topologies is compared, and a sensitivity analysis on the parameters of reservoir topologies is conducted. A case study is given to show the excellent performance of ESNs with the developed simplified reservoir topologies in energy consumption prediction of an office building.

However, several important factors may affect the energy consumption in an office building, including the weather conditions, building envelope, internal loads, outdoor air flow rate and infiltration rate, etc. Only the available data of outside air temperature and building occupancy are presented in this paper. In future work, we will try to collect and incorporate more input data into the prediction models, with a view to further improving the prediction accuracy. On the other hand, we focus on the energy consumption in working rooms in this paper, but we will collect the data of energy consumption in public areas in the future, including the passages, restrooms, etc., so that the entire energy consumption of the building can be predicted.

In addition, based on the prediction results of energy consumption in the office building, a control algorithm may be developed to optimize the energy consumption in the office building by installing an energy storage equipment in each room as the control variable. Moreover, renewable resources such as solar and wind energies may be introduced into the control system for further optimization of energy consumption.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grants 61273140, 61374105, 61503377, 61503379, 61533017, and U1501251.

References

[1] Y. Wu, X. Zhao, K. Li, M. Zheng, S. Li, Energy saving–Another perspective for parameter optimization of P and PI controllers, Neurocomputing 174 (2016) 500–513.
[2] S. Arslan, The influence of environment education on critical thinking and environmental attitude, Procedia—Social Behav. Sci. 55 (2012) 902–909.
[3] D. Streimikiene, Residential energy consumption trends, main drivers and policies in Lithuania, Renew. Sustain. Energy Rev. 35 (2014) 285–293.
[4] M. Santamouris, Energy and Climate in the Urban Built Environment, Routledge, London, 2013.
[5] H. Chen, W. Lee, X. Wang, Energy assessment of office buildings in China using China building energy codes and LEED 2.2, Energy Build. 86 (2015) 514–524.
[6] A. Ahmad, M. Hassan, M. Abdullah, H. Rahman, F. Hussin, H. Abdullah, R. Saidur, A review on applications of ANN and SVM for building electrical energy consumption forecasting, Renew. Sustain. Energy Rev. 33 (2014) 102–109.
[7] H. Zhao, F. Magouls, A review on the prediction of building energy consumption, Renew. Sustain. Energy Rev. 16 (6) (2012) 3586–3592.
[8] Y. Pan, Z. Huang, G. Wu, Calibrated building energy simulation and its application in a high-rise commercial building in Shanghai, Energy Build. 39 (2007) 651–657.
[9] C. Fan, F. Xiao, S. Wang, Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques, Appl. Energy 127 (2014) 1–10.
[10] D. Srinivasan, Energy demand prediction using GMDH networks, Neurocomputing 72 (2008) 625–629.

[11] D. Hsu, Comparison of integrated clustering methods for accurate and stable prediction of building energy consumption data, Appl. Energy 160 (2015) 153–163.
[12] Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida, Applying support vector machine to predict hourly cooling load in the building, Appl. Energy 86 (10) (2009) 2249–2256.
[13] L. Ciabattoni, F. Ferracuti, M. Grisostomi, G. Ippoliti, S. Longhi, Fuzzy logic based economical analysis of photovoltaic energy management, Neurocomputing 170 (2015) 296–305.
[14] C. Peng, Y. Bai, X. Gong, Q. Gao, C. Zhao, Y. Tian, Modeling and robust backstepping sliding mode control with adaptive RBFNN for a novel coaxial eight-rotor UAV, IEEE/CAA J. Autom. Sin. 2 (1) (2015) 56–64.
[15] J. Zhang, C. Sun, R. Zhang, C. Qian, Adaptive sliding mode control for re-entry attitude of near space hypersonic vehicle based on backstepping design, IEEE/CAA J. Autom. Sin. 2 (1) (2015) 94–101.
[16] S. Wang, Y. Zhang, Y. Jin, Y. Zhang, Neural control of hypersonic flight dynamics with actuator fault and constraint, Sci. China Inf. Sci. 58 (2015).
[17] Q. Wei, D. Liu, A novel policy iteration based deterministic Q-learning for discrete-time nonlinear systems, Sci. China Inf. Sci. 58 (2015).
[18] A. Cherif, H. Cardot, R. Bone, SOM time series clustering and prediction with recurrent neural networks, Neurocomputing 74 (2011) 1936–1944.
[19] R. Chandra, M. Zhang, Cooperative coevolution of Elman recurrent neural networks for chaotic time series prediction, Neurocomputing 86 (2012) 116–123.
[20] X. Wang, L. Ma, B. Wang, T. Wang, A hybrid optimization-based recurrent neural network for real-time data prediction, Neurocomputing 120 (2013) 547–559.
[21] C. Smith, Y. Jin, Evolutionary multi-objective generation of recurrent neural network ensembles for time series prediction, Neurocomputing 143 (2014) 302–311.
[22] H. Chitsaz, H. Shaker, H. Zareipour, D. Wood, N. Amjady, Short-term electricity load forecasting of buildings in microgrids, Energy Build. 99 (2015) 50–60.
[23] S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 1998.
[24] H. Jaeger, The 'Echo State' Approach to Analysing and Training Recurrent Neural Networks, GMD Technical Report 148, German National Research Center for Information Technology, 2001.
[25] H. Jaeger, H. Haas, Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication, Science 304 (5667) (2004) 78–80.
[26] C. Gallicchio, A. Micheli, Tree Echo State Networks, Neurocomputing 101 (2013) 319–337.
[27] Q.-L. Man, W.-B. Chen, Modular state space of echo state network, Neurocomputing 122 (2013) 406–417.
[28] L. Boccato, R. Attux, F.J.V. Zuben, Self-organization and lateral interaction in echo state network reservoirs, Neurocomputing 138 (2014) 297–309.
[29] H. Wang, X. Yan, Improved simple deterministically constructed Cycle Reservoir Network with Sensitive Iterative Pruning Algorithm, Neurocomputing 145 (2014) 353–362.
[30] S. Yuenyong, A. Nishihara, Evolutionary pre-training for CRJ-type reservoir of echo state networks, Neurocomputing 149 (2015) 1324–1329.
[31] S. Otte, M.V. Butz, D. Koryakin, F. Becker, M. Liwicki, A. Zell, Optimizing recurrent reservoirs with neuro-evolution, Neurocomputing 192 (2016) 128–138.
[32] B. Liebald, Exploration of effects of different network topologies on the ESN signal crosscorrelation matrix spectrum, International University Bremen, 2004.
[33] Z. Deng, Y. Zhang, Collective behavior of a small-world recurrent neural system with scale-free distribution, IEEE Trans. Neural Netw. 18 (5) (2007) 1364–1375.
[34] Q. Song, Z. Feng, Effects of connectivity structure of complex echo state network on its prediction performance for nonlinear time series, Neurocomputing 73 (10–12) (2010) 2177–2185.
[35] A. Rodan, P. Tino, Minimum complexity echo state network, IEEE Trans. Neural Netw. 22 (1) (2011) 131–144.
[36] H. Jaeger, M. Lukoevicius, D. Popovici, U. Siewert, Optimization and applications of echo state networks with leaky-integrator neurons, Neural Netw. 20 (3) (2007) 335–352.
[37] J.J. Steil, Online reservoir adaptation by intrinsic plasticity for backpropagation-decorrelation and echo state learning, Neural Netw. 20 (3) (2007) 353–364.
[38] J. Triesch, A gradient rule for the plasticity of a neuron's intrinsic excitability, in: 15th International Conference on Artificial Neural Networks, 2005, pp. 65–70.
[39] B. Schrauwen, M. Wardermann, D. Verstraeten, J.J. Steil, D. Stroobandt, Improving reservoirs using intrinsic plasticity, Neurocomputing 71 (7–9) (2008) 1159–1171.
[40] S.-X. Lun, X.-S. Yao, H.-Y. Qi, H.-F. Hu, A novel model of leaky integrator echo state network for time-series prediction, Neurocomputing 159 (2015) 58–66.
[41] C. Sheng, J. Zhao, Y. Liu, W. Wang, Prediction for noisy nonlinear time series by echo state network based on dual estimation, Neurocomputing 82 (2012) 186–195.
[42] M. Han, M. Xu, X. Liu, X. Wang, Online multivariate time series prediction using SCKF-γ ESN model, Neurocomputing 147 (2015) 315–323.
[43] M.C. Ozturk, J.C. Principe, An associative memory readout for ESNs with applications to dynamical pattern recognition, Neural Netw. 20 (3) (2007) 377–390.
[44] M.D. Skowronski, J.G. Harris, Noise-robust automatic speech recognition using a predictive echo state network, IEEE Trans. Audio Speech Lang. Process. 15 (5) (2007) 1724–1730.
[45] Y. Xia, B. Jelfs, Marc M. Van Hulle, J.C. Principe, D.P. Mandic, An augmented echo state network for nonlinear adaptive filtering of complex noncircular signals, IEEE Trans. Neural Netw. 22 (1) (2011) 74–83.
[46] F.M. Bianchi, E. De Santis, A. Rizzi, A. Sadeghian, Short-term electric load forecasting using echo state networks and PCA decomposition, IEEE Access 3 (2015) 1931–1943.
[47] F.M. Bianchi, S. Scardapane, A. Uncini, A. Rizzi, A. Sadeghian, Prediction of telephone calls load using Echo State Network with exogenous variables, Neural Netw. 71 (2015) 204–213.
[48] A. Deihimi, H. Showkati, Application of echo state networks in short-term electric load forecasting, Energy 39 (1) (2012) 327–340.
[49] E. Crisostomi, C. Gallicchio, A. Micheli, M. Raugi, M. Tucci, Prediction of the Italian electricity price for smart grid applications, Neurocomputing 170 (2015) 286–295.
[50] G. Manjunath, H. Jaeger, Echo state property linked to an input: Exploring a fundamental characteristic of recurrent neural networks, Neural Comput. 25 (3) (2013) 671–696.
[51] B. Zhang, D.J. Miller, Y. Wang, Nonlinear system modeling with random matrices: Echo state networks revisited, IEEE Trans. Neural Netw. Learn. Syst. 23 (1) (2012) 175–182.
[52] D. Serre, Matrices: Theory and Applications, Springer, New York, 2002.
[53] C.R. Rao, S.K. Mitra, Generalized Inverse of Matrices and its Applications, Wiley, New York, 1971.

Guang Shi received the B.S. degree in automation from Zhejiang University, Hangzhou, China, in July, 2012. Currently, he is working towards the Ph.D. degree at The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. His research interests include neural networks, adaptive dynamic programming, optimal control and energy management in smart grids.

Derong Liu received the Ph.D. degree in electrical engineering from the University of Notre Dame, Notre Dame, IN, USA, in 1994. From 1993 to 1995, he was a Staff Fellow with the General Motors Research and Development Center, Warren, MI, USA. From 1995 to 1999, he was an Assistant Professor with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, USA. In 1999, he joined the University of Illinois at Chicago, Chicago, IL, USA, where he became a Full Professor of electrical and computer engineering and of computer science in 2006. He is the author of 17 books.

Dr. Liu currently serves as the Editor-in-Chief of Artificial Intelligence Review. He received the Faculty Early Career Development Award from the National Science Foundation in 1999, the University Scholar Award from the University of Illinois in 2006, and the Overseas Outstanding Young Scholar Award from the National Natural Science Foundation of China in 2008. He was selected for the "100 Talents Program" by the Chinese Academy of Sciences in 2008.

Qinglai Wei received the B.S. degree in automation, the M.S. degree in control theory and control engineering, and the Ph.D. degree in control theory and control engineering, from the Northeastern University, Shenyang, China, in 2002, 2005, and 2008, respectively. From 2009 to 2011, he was a postdoctoral fellow with State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. He is currently an Assistant Research Fellow of the institute. His research interests include neural-networks-based control, adaptive dynamic programming, optimal control, nonlinear system and their industrial applications.