Neurocomputing 432 (2021) 32–43

Contents lists available at ScienceDirect

Neurocomputing

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

Synergies between synaptic and intrinsic plasticity in echo state networks

Xinjie Wang [a,b], Yaochu Jin [a,b,c,*], Kuangrong Hao [b]

[a] Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China
[b] Engineering Research Center of Digitized Textile and Apparel Technology, Ministry of Education, College of Information Science and Technology, Donghua University, Shanghai 201620, China
[c] Nature Inspired Computing and Engineering, Department of Computer Science, University of Surrey, UK

ARTICLE INFO

Article history:
Received 1 December 2019
Revised 6 October 2020
Accepted 5 December 2020
Available online 16 December 2020
Communicated by Steven Hoi

Keywords:
Echo state networks
Synaptic plasticity
Intrinsic plasticity
Synergistic learning
Regression
Classification

ABSTRACT

Synaptic plasticity and intrinsic plasticity, as two of the most common neural plasticity mechanisms, occur in all neural circuits throughout life. Neurobiological studies indicated that the interplay between synaptic and intrinsic plasticity contributes to the adaptation of the nervous system to different synaptic input signals. However, most existing computational models of neural plasticity consider these two plasticity mechanisms separately, which is biologically implausible. In this paper, a synergistic plasticity learning rule is proposed to adapt the reservoir connections in echo state networks (ESNs), which not only takes into account the regulation of synaptic weights, but also considers the adjustment of neuronal intrinsic excitability. The proposed synergetic plasticity rule is verified on a number of prediction and classification benchmark problems and our empirical results demonstrate that the ESN with synergistic plasticity learning rule performs much better than the state-of-the-art ESN models, and an ESN with a single neural plasticity rule.

1. Introduction

Reservoir computing is a new class of recurrent neural networks with a large-scale sparsely connected state-space structure (the reservoir) and an adaptable readout [1], which can effectively avoid the vanishing gradient problem and considerably reduce the computational complexity compared with the traditional recurrent neural networks (RNNs) [2–4]. Over the past two decades, echo state networks (ESNs) [5] and liquid state machines (LSMs) [6,7], both inspired by neurophysiological experiments, have been recognized as the two most efficient and common reservoir computing models for training RNNs. LSMs are commonly used to perform real-time computing on temporal signals by adapting the dynamics of a set of coupled spiking neuron models [6]. This paper focuses on the ESN model.

In recent years, the ESN and its variants have been widely applied to various synthetic datasets and real-life applications, such as time series prediction and classification [8,9], electroencephalogram (EEG) feature extraction and emotion recognition [10,11], and human action recognition [12].

The original ESN model can be considered a three-layer neural network model, with an input layer having fixed connections, a sparsely connected hidden layer, also referred to as the reservoir, and a readout layer [13]. The original ESN only needs to train the weights of the reservoir-to-readout connections using a simple linear classification or regression learning algorithm. However, such a simple reservoir computing model also faces some challenges in accuracy and robustness [14]. Specifically, the randomly generated reservoir weights and other neuronal parameters may result in instability in learning and yield a high variance in learning performance. In addition, the randomly generated and fixed weights of the reservoir also limit the learning performance of the ESN.

To address the above-mentioned weaknesses of the original ESN, many novel methods have been proposed to construct more effective reservoir topologies or generate the weight matrices adaptively. For instance, Jaeger et al. [15] proposed a leaky-integrator ESN (LESN) that simply modified the reservoir dynamic equations of the initial ESN to enhance the flexibility of the network. Yang et al. [16] presented a polynomial ESN (PESN) by employing the polynomial functions of input variables into output weights to improve the prediction performance. Bianchi et al. [17] investigated the dynamics of the reservoir with recurrence plots

* Corresponding author at: Department of Computer Science, Universoty of Surrey, Guildford, GU2 7XH, United Kingdom.
E-mail addresses: wangxinjie0621@foxmail.com (X. Wang), yaochu.jin@surrey.ac.uk (Y. Jin), krhao@dhu.edu.cn (K. Hao).

https://doi.org/10.1016/j.neucom.2020.12.007
0925-2312/

and recurrence quantification analysis to quantify network stability. Besides, simple cycle reservoir (SCR) topology was proposed by Rodan and Tino [1], which can minimize the computational complexity in constructing the reservoir without deteriorating the learning performance. Kawai et al. [18] constructed a small-world echo state network, which exhibited the best learning performance when the spectral radius was more than one. That is, the small-world network provided the ESN with a stable echo state property over a broad range of the weight matrix. Similarly, Kudithipudi et al. [19] proposed a neuromemristive reservoir structure with doubly twisted toroidal structure to reduce power consumption. Furthermore, some ESNs with multiple reservoirs have been proposed to accomplish specific complex tasks [4,20,21]. Ma et al. [20] constructed a convolutional multitimescale ESN with multiple reservoirs and a convolutional layer for capturing multiple time scale structures of temporal data and extracting temporal dependencies of different timescale. A growing ESN (GESN) with multiple sub-reservoirs is proposed to automatically design ESN [21]. Xu et al. [22] presented a spatio-temporal prediction model based on ESNs, in which the elastic-net algorithm and the ESN are used to capture the optimal spatial coefficients and the dynamic characteristics of the series, respectively.

In addition, various optimization techniques have been presented for optimizing the hyperparameters and developing appropriate topologies of the ESN, such as the gradient based method [23], Bayesian optimization algorithm [24], neuroevolution of augmented topologies [25], and single- and multi-objective optimization algorithms [26].

As an unsupervised biologically inspired adaptation mechanism, neural plasticity is a novel way to improve the learning performance [2,27]. So far, two neural plasticity mechanisms, synaptic and intrinsic plasticity, have most widely been studied. In fact, most theoretical and experimental research on the neural plasticity has focused on regulations of synaptic strength or connectivity for keeping a stable learning performance [28]. Synaptic plasticity, as the most common and important neural plasticity mechanism, refers to the activity-dependent modifications of the synaptic strength or efficacy [29,30]. It can contribute to incorporating transient experiences into persistent memory traces [31]. The change in the synaptic strength or efficacy can be caused by either the firing rate of the presynaptic neuron or by the amount of neurotransmitter released [30,28]. Several computational models for neural plasticity have been presented, including the Hebbian rule [32], the BCM rule [33], the anti-Oja rule [34], and the spike-timing-dependent plasticity [35], among many others. Apart from the synaptic plasticity, some studies later found that an individual biological neuron can also alter its intrinsic excitability to adapt to different synaptic inputs [36,37]. The neurons achieve this by adjusting the positions and the slopes of the neuronal responses to help neurons preserve the output characteristics for different synaptic inputs and maintain a suitable firing rate level [38]. This neural mechanism is known as intrinsic plasticity.

In fact, the goal of intrinsic plasticity is to maximize the information entropy of output neurons while obeying constraints on the mean of the output distribution, which can contribute to improving the stability of ESNs [2,39]. In addition, Koprinkova-Hristova also pointed out that intrinsic plasticity can capture the structure of input data and represent it in the form of the reservoir steady state that could be useful for data clustering capability [40,41]. Some neurobiological experimental phenomena further indicated that the responses of mammalian visual cortical neurons to natural scenes are approximately exponential spike-count distributions [42] that have the highest information entropy among all non-negative distributions with a same mean [37]. Meanwhile, some learning rules have been proposed to show that how adjustment in neuronal excitability can be achieved with ion channel based realistic network model to adapt to changes in the synaptic input [43].

In artificial neural networks, computational models of intrinsic plasticity have been examined to see if they can improve the learning performance. Experiments in [37] confirm that intrinsic plasticity plays an important role in shaping neuron's activation function to maintain neuron's homeostasis of ultimate outputs. Specifically, existing intrinsic plasticity rules generally assumed that the single neuronal firing rate obeys a desired probability distribution with a constant mean. Then, the plasticity rule can be derived by minimizing the Kullback-Leiber divergence between the firing rate distribution and the desired distribution. In previous studies, different types desired distributions were used to appropriately describe the real distribution of biological neurons, including exponential-based [37], Gaussian-based [11], and Weibull-based [38,44].

In biological nervous systems, the interplay between the synaptic inputs and the neuronal intrinsic excitability controls the response of the neuron and neurotransmitter release [45]. Neural plasticity has long been considered the neuronal basis of learning and memory, which can result from alterations in synaptic strength or connectivity, and from changes in neuronal excitability [28].

Synaptic and intrinsic plasticity, as two most common plasticity mechanisms, have been widely studied. Several recent biological experiments further indicated that the summed impact of multiple forms of synaptic and intrinsic homeostatic mechanisms contribute to adaptive changes to brain function and behavior [30]. Lambo and Turrigiano [46] tested the effects of monocular deprivation of excitability of the mouse visual cortex and found that intrinsic excitability of pyramidal neurons increases with the increase of the excitatory postsynaptic current amplitude. Meanwhile, this homeostatic increase of intrinsic excitability amplifies the changes in excitatory postsynaptic strength [46]. A similar phenomenon found that action potential firing in the vestibular system results from intrinsic properties of vestibular nucleus neurons and excitatory synapses from vestibular nerve axons [47]. Electrophysiological recordings from the vestibular system also found that synaptic and intrinsic plasticity can conduce to the restoration of the neuronal activity in vestibular nuclei after vestibular dysfunction [48]. In addition, neuro- modulation and some plasticity-related proteins could also influence synaptic efficacy and neuronal intrinsic excitability, which plays important roles in memory formation and maintenance [28,49].

In addition to these theoretical and experimental findings in vivo, a few studies attempt to construct synaptic and intrinsic plasticity cooperation mechanisms. Triesch [50] pointed out that the interaction of intrinsic and synaptic plasticity mechanisms can discover heavy-tailed directions in the input space for a simple single sigmoid neuron. Li et al. [51] presented a synergistic learning model that adopts the error-entropy minimization algorithm to update the weight matrices and an information-maximization algorithm to adjust the activation function. However, the error-entropy minimization and the information-maximization algorithms are not biologically inspired neural plasticity learning rules.

The following observations can be made on neural plasticity based on the above studies. First, both synaptic plasticity and intrinsic plasticity occur in all neural circuits throughout life. Second, extensive neurobiological research indicated that the interplay between synaptic and intrinsic plasticity contributes to the adaption of the nervous system to changes in internal and external environments. Third, existing computational models of neural plasticity proposed so far have barely simultaneously considered the synergetic effects of synaptic and intrinsic plasticity, which is biologically implausible.

Thus, the main motivation of this paper is to bridge the gap between biological plasticity mechanisms and computational

models. Based on the biological findings, we propose a synergistic plasticity learning rule that takes into account both the regulation of synaptic weights using the synaptic plasticity and the adjustment of neuronal intrinsic excitability using the intrinsic plasticity, thereby further improving the learning performance of the ESN. The effectiveness of the proposed synergistic plasticity model is verified on several regression and classification problems, including the tenth-order nonlinear autoregressive moving average (NARMA) system, the Henon map system, the Lorenz system, and the EEG eye state classification task.

The rest of this paper is organized as follows. In Section 2, the ESN model, the synaptic rule, and the intrinsic plasticity rule are briefly introduced. The proposed synergistic plasticity learning rule is detailed in Section 3. Experimental settings and results are presented in Sections 4 and 5, respectively. Conclusions and future work are given in Section 6.

2. Related work

2.1. The ESN architecture

The ESN is a particular paradigm of recurrent neural networks. The standard ESN model without output feedback contains an input layer, a reservoir and an output layer. A large-scale sparsely randomly connected reservoir is usually used for building a high-dimensional non-linear mapping of the input signals to the output. In the canonical ESN, the weights between the neurons in the reservoir are randomly generated and kept constant and only the output weights are trained [52]. The basic structure of ESN is shown in Fig. 1.

Consider an ESN model with K input units, M reservoir units and L output units. In the canonical ESN, the weights of the input-to-reservoir connections W^{in} and those of the internal reservoir W^{res} are randomly generated and only the connection weights of reservoir-to-readout layer W^{out} will be trained. The equations of the input state $x^{in}(t)$ and the activation state $x(t)$ of the reservoir at time t are expressed as follows [53]:

$$x^{in}(t) = W^{in}u(t) + W^{res}x(t-1), \tag{1}$$

$$x(t) = f(a_i(t)x^{in}(t) + b_i(t)), \tag{2}$$

where, $u(t)$ is the external input at time t, and f is the sigmoid function. The coefficients $a_i(t)$ and $b_i(t)$ are the gain and bias of the ith reservoir neuron's activation function at time t, respectively. Normally the gain and the bias are set to 1 and 0, respectively. The readout $y(t)$ at time t is calculated by:

$$y(t) = W^{out}x(t) \tag{3}$$

The normalized root mean square error (*NRMSE*) and the softmax function are adopted as the prediction error of regression problems and recognition accuracy of classification tasks, respectively, which are given in Eqs. (4) and (5):

$$NRMSE(W^{out}) = \sqrt{\frac{\sum_{t=1}^{T}(y^{desired}(t) - y(t))^2}{T\sigma^2}}, \tag{4}$$

$$P(i) = \text{softmax}(y_i(t)) = \frac{\exp(y_i(t))}{\sum_{l=1}^{L}\exp(y_i(t))} \tag{5}$$

where, σ^2 represents the variance of the desired outputs in the *NRMSE* function. For classification task, the softmax function is adopted for calculating the classification accuracy. The number of the input layer neurons and output layer neurons depends on the number of attributes and classes. The output node with the highest probability and the corresponding class is used as the actual class label. The classification accuracy can be calculated by comparing the desired class label and the actual class label. In the softmax function, $y_i(t)$ is the actual output of the readout neuron i at time t, L represents the number of labels on samples, $P(i)$ is the probability at which the sample belongs to the ith label.

The readout weights W^{out} can be computed by minimizing the difference between the actual output y and the desired output $y^{desired}$ of the network. The reservoir activation state and readout vector in each moment in the training phase will be stored in X and Y, respectively. The readout weights are calculated by:

$$W^{out} = (X^{\mathrm{T}}X)^{-1}X \cdot Y, \tag{6}$$

2.2. Synaptic plasticity

A large number of studies on neural plasticity revealed that synaptic plasticity can contribute to stabilizing the synaptic connections and adapting to complex internal and external environments [54]. It has also been demonstrated that most of synaptic strengths and connectivity are altered by the temporal pattern of action potential firing of the presynaptic and postsynaptic neurons [28,49].

Most of existing synaptic plasticity learning rules are based on the Hebbian rule [32,29], which can adjust the connection weights through the output states of the neurons between presynaptic and postsynaptic. The Hebbian learning rule is defined as:

$$\Delta W_{ji}^{res}(t) = \eta x_i(t)x_j(t), \tag{7}$$

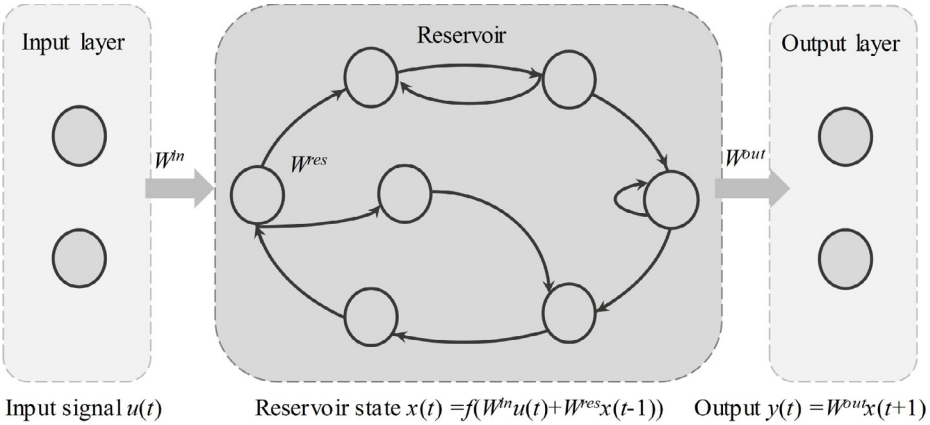

Fig. 1. The basic structure of the ESN model.

where $\Delta W_{ji}^{res}(t)$ represents the weight change between the presynaptic neuron i and postsynaptic neuron j at time t. x_i and x_j are the outputs of neuron i and neuron j, respectively. η is a small positive learning rate parameter. In fact, the Hebbian learning rule is inherently unstable. Connection weights will increase unrealistically, since all directions of modifications are positive [55]. An opposite learning rule, called anti- Hebbian rule, is presented, which tries to stabilize the Hebbian rule by inhibiting joint firing of many neurons [56]. In our previous studies, the roles of Hebbian rule and anti-Hebbian were discussed in detail [57].

A saturation or normalization method can solve the instability problem. Oja proposed a modified rule based on the original Hebbian rule by normalizing all weights connected to postsynaptic neuron [58]:

$$\sqrt{\sum_{i=1}^{n} W_{ji}^{res}(t)^2} = 1, \tag{8}$$

where W_{ji}^{res} is the synaptic weight between the presynaptic neuron i and postsynaptic neuron j. All synaptic weights need to be normalized in the following form:

$$W_{ji}^{res}(t+1) = \frac{W_{ji}^{res}(t) + \eta x_i(t) x_j(t)}{\sqrt{\sum_{i=1}^{n} [W_{ji}^{res}(t) + \eta x_i(t) x_j(t)]^2}} \tag{9}$$

When the learning rate η goes to 0, an equivalent form of Eq. (9) is again obtained, which can be expanded as a power series [58]. This learning rule with a forgetting term is called the Oja learning rule:

$$W_{ji}^{res}(t+1) = W_{ji}^{res}(t) + \eta x_j(t)[x_i(t) - x_j(t) W_{ji}^{res}(t)], \tag{10}$$

where $x_j(t)^2 W_{ji}^{res}(t)$ is the weight forgetting term. The anti-Oja learning rule with a negative weight update can be regarded as an inverted form of the Oja learning rule.

$$W_{ji}^{res}(t+1) = W_{ji}^{res}(t) - \eta x_j(t)[x_i(t) - x_j(t) W_{ji}^{res}(t)] \tag{11}$$

In addition, different computational models of synaptic plasticity are presented to preserve the stability of the regulation process, including the BCM rule [33], synaptic scaling [59] and metaplasticity [60]. In this work, the anti-Oja learning rule is used to adjust the connection weights between the neurons in the reservoir.

2.3. Intrinsic plasticity

Many studies on neural plasticity in both neural circuits and computational modeling have focused on synaptic plasticity. However, intrinsic plasticity also plays crucial roles in changing the neuronal excitability, which can be affected by the distribution and number of ion channels [28].

In addition, biological experiments indicate that neuronal firing rate distributions in mammalian primary and inferior temporal visual cortical areas in response to natural scenes are approximately exponential spike-count distributions [42]. Actually, these observed approximately exponential distributions almost have the maximum information among all non-negative distributions with a fixed mean [42,38]. Thus, exponential spike-count distributions are believed to be able to maintain the neuronal firing level in a desirable region and maximize information transmission to the downstream targets [38,50].

Based on the above experimental observations, Triesch [37] presented a computational model of intrinsic plasticity by matching the neuronal actual output distribution to a desired exponential distribution with a fixed mean. Then, based on the information maximization theory, this intrinsic plasticity rule can be derived by minimizing the Kullback-Leiber divergence between the actual output distribution and the desired distribution. In this work, a sigmoid neuron is used to activation function:

$$x = f_{a,b}(x^{in}) = \frac{1}{1 + \exp(-(ax^{in} + b))} \tag{12}$$

Specifically, the intrinsic plasticity rule based on information theory adjusts the gain a and the bias b of the neuron to make actual output distribution follow an exponential distribution with a fixed mean. Two parameters of the neuron are updated by the stochastic gradient descent rule in each iteration as shown in the following [37]:

$$a_i(t+1) = a_i(t) + \frac{\eta_{IP}}{a_i(t)} + x_i^{in}(t)\Delta b_i(t) \tag{13}$$

$$b_i(t+1) = b_i(t) + \eta_{IP}(1 - (2 + \frac{1}{\mu})x_i(t) + \frac{1}{\mu}x_i^2(t)) \tag{14}$$

The parameters $\Delta a_i(t+1)$ and $\Delta b_i(t+1)$ are the changes of the gain and the bias of the i-th reservoir neuronal activation function at the time $t+1$, respectively. μ represents the mean of the desired distribution. η_{IP} is the learning rate of the intrinsic plasticity rule.

3. Synergistic plasticity learning rule

Based on the theoretical and experimental findings on the neural plasticity in nervous systems, here we propose a synergistic plasticity rule that simultaneously take into account the synaptic and intrinsic plasticity learning rule. Specifically, the anti-Oja learning rule and the exponential distribution- based intrinsic plasticity are adopted to regulate the synaptic weights and to adjust the neuron's activation function, respectively. A diagram of the regulation processes of the ESN with the synergistic learning rule is shown in Fig. 2.

As shown in Fig. 2, the weights within the reservoir can be adjusted through the output states of the neurons they are connected. The gain and bias of activation function can also be modulated by the input and activation states. That is, changes in both synaptic strength and neuronal intrinsic excitability affect directly the ultimate outputs of the neuron, which also coincide well with neurobiological findings [28,30,45]. The synergistic learning rule is defined as:

$$W_{ji}^{res}(t+1) = W_{ji}^{res}(t) - \eta x_j(t)[x_i(t) - x_j(t) W_{ji}^{res}(t)] \tag{15}$$

$$a_i(t+1) = a_i(t) + \frac{\eta_{IP}}{a_i(t)} + x_i^{in}(t)\Delta b_i(t) \tag{16}$$

$$b_i(t+1) = b_i(t) + \eta_{IP}(1 - (2 + \frac{1}{\mu})x_i(t) + \frac{1}{\mu}x_i^2(t)) \tag{17}$$

The main difference between the proposed synergistic plasticity rule and existing plasticity rules is that the synergistic plasticity rule combines synaptic plasticity with intrinsic plasticity, as shown in the above equations. As a result, the interplay between the synaptic and intrinsic plasticity is taken into account.

4. Experimental setup

In the following we give details about the experimental setup. Specifically, we describe the empirical parameter setting, the four benchmark problems and the algorithms under comparison.

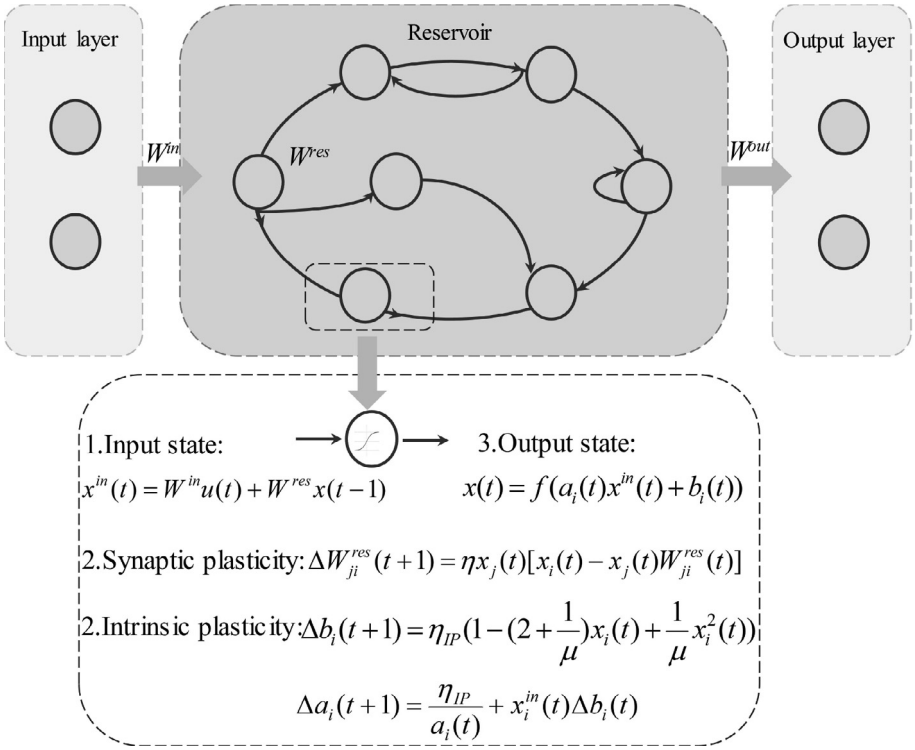

Fig. 2. Regulation processes of the ESN with synergistic learning rule.

4.1. Benchmark problems

4.1.1. Prediction of the NARMA system

The NARMA system is a nonlinear system identification task, which has been widely used as a benchmark problem of time series prediction [1]. The chaotic characteristics and the long memory of the NARMA system make the forecast of this system difficult [61]. In this work, a tenth-order NARMA system is used to evaluate the learning performance of the ESN. The dynamics of tenth-order NARMA system is described by:

$$x(n+1) = 0.3x(n) + 0.05x(n)\sum_{i=0}^{9}x(n-i) + 1.5u(n-9)u(n) + 0.1 \tag{18}$$

where, the input sequence $u(n)$ is generated from a uniform distribution in the range [0, 0.5]. $x(n)$ represents the system output at time n. The initial values of the NARMA system $x(1), x(2), \ldots, x(10)$ are set to 0. In this task, 5200 samples are generated, of which the first 1000 training samples are used to washout the initial transient of the network, and the following 1800 training data are employed to adjust the gain and the bias of activation function and connection weights within the reservoir. The rest 2400 samples are used for test.

4.1.2. Prediction of the Henon map system

The Henon map system is a typical nonlinear discrete-time dynamical system, which was first proposed by Henon [62]. Due to the chaotic characteristics of the system, the Henon map has been widely used in image encryption [63] and chaos control [64], which is described by:

$$\begin{cases} u_1(t+1) = u_2(t) - au_1^2(t) + 1 \\ u_2(t+1) = bu_1(t) \end{cases} \tag{19}$$

This dynamical system displays chaotic phenomenon, when a = 1.4 and b = 0.3. This task is to forecast $u_1(t+10)$ using $u_1(t)$ and $u_2(t)$, when the initial inputs equal zero. 1250 samples are generated from the Henon map system, of which the first 200 training samples are used to washout the network, and the following 800 training data are employed to modulate neuronal activation function and connection weights within the reservoir. All the rest samples are used as the test dataset.

4.1.3. Prediction of the Lorenz system

The Lorenz system is also a classical benchmark prediction task and the discrete-time Lorenz system is described by [23,26,65]:

$$\begin{cases} \frac{x(t+1)-x(t)}{dt} = a(-x(t) + y(t)) \\ \frac{y(t+1)-y(t)}{dt} = bx(t) - y(t) - x(t)z(t) \\ \frac{z(t+1)-z(t)}{dt} = x(t)y(t) - cz(t) \end{cases} \tag{20}$$

where t is the time instant. In general, the system constants are set to a = 10, b = 28, c = 8/3. This task is to predict $x(t+10)$ based on historical data $x(t-20), x(t-10)$ and $x(t)$. The initial condition of the system is $x(0)$=$y(0)$=1, $z(0)$=0, and the time interval $d(t)$ is 0.01. 3500 samples are generated the fourth order Runge–Kutta method. The first 500 training samples are used to washout, the next 2000 samples for training, and the last 1000 samples are used for test.

4.1.4. Classification of the EEG eye state

The EEG eye state dataset is employed to verify the proposed synergistic plasticity learning rule for classification. This dataset was taken from the Machine Learning Repository, University of California Irvine [66]. The dataset contains 14,980 samples, each having 14 input features and two classes, and the features were recorded from 14 different electrodes with the Emotiv EEG Neuroheadset, namely AF3, T7, P8, F8, F7, P7, T8, AF4, F3, O1, FC6, FC5, O2, F4 [67]. Eye-closed and eye-open state were captured via a camera during the EEG measurement. The ratio of training to test set is 4:1.

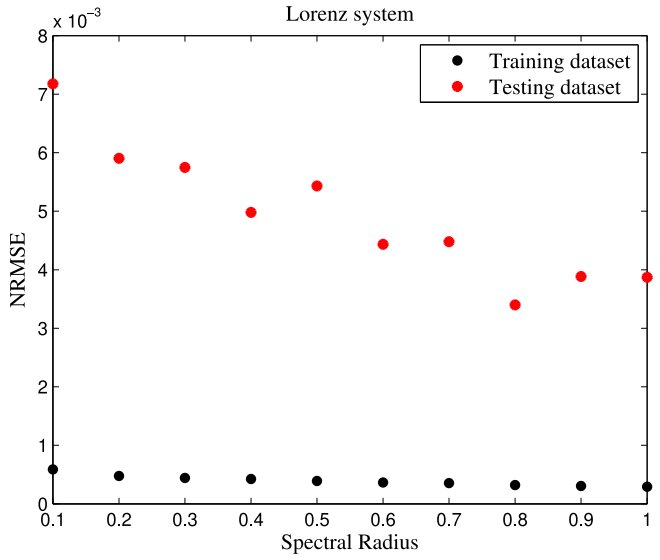

Fig. 3. Effects of spectral radius on the learning performance on the Lorenz system.

4.2. Empirical parameter setting

In this part, pilot studies are carried out to analyze the effects of the connection weights and spectral radius, the reservoir size, and the parameters in the plasticity rules on the learning performance.

4.2.1. The setting of the connection weights and spectral radius

For a standard three layers ESN, the initial connection weights of input-to-reservoir W^{in} and internal reservoir W_0^{res} are randomly generated within the interval of [-1, 1]. In general, W^{in} also need to be scaled in such a way that the activation function works in the linear region, which is between −0.1 and 0.1. To ensure the

Table 1
Learning rates of synaptic and intrinsic plasticity rules.

Dataset	Synaptic plasticity	Intrinsic plasticity
EEG eye state	0.000001	0.00001
NARMA system	0.00001	0.00001
Henon map system	0.000001	0.00001
Lorenz system	0.00001	0.0001

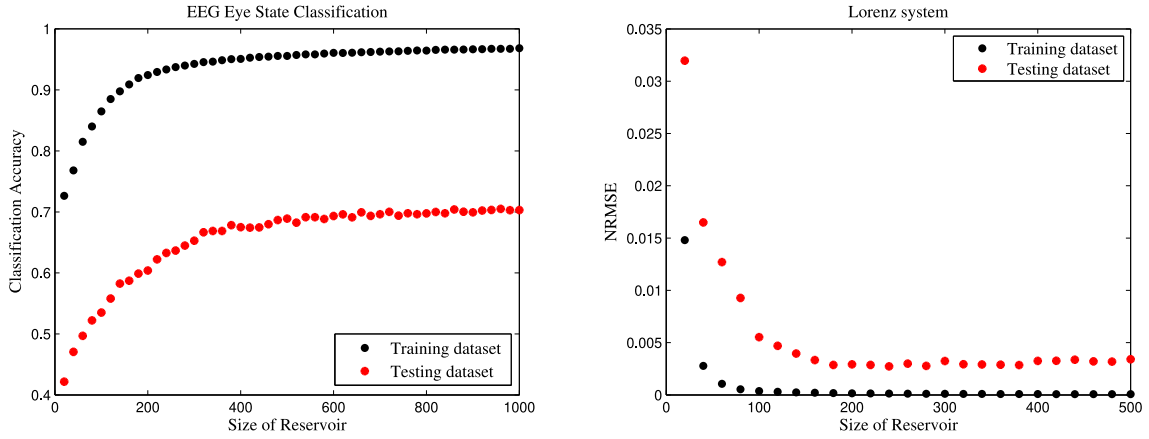

Fig. 4. Effects of reservoir size on the learning performance on the EEG eye state dataset and the Lorenz system.

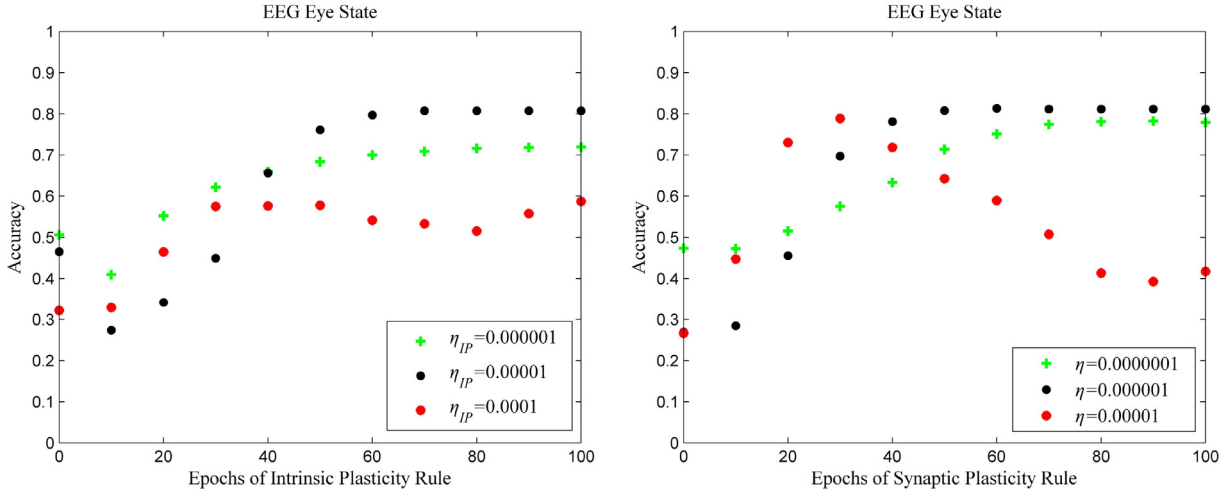

Fig. 5. Convergence profiles for different learning rates of the synaptic plasticity and intrinsic plasticity on the EEG eye state dataset.

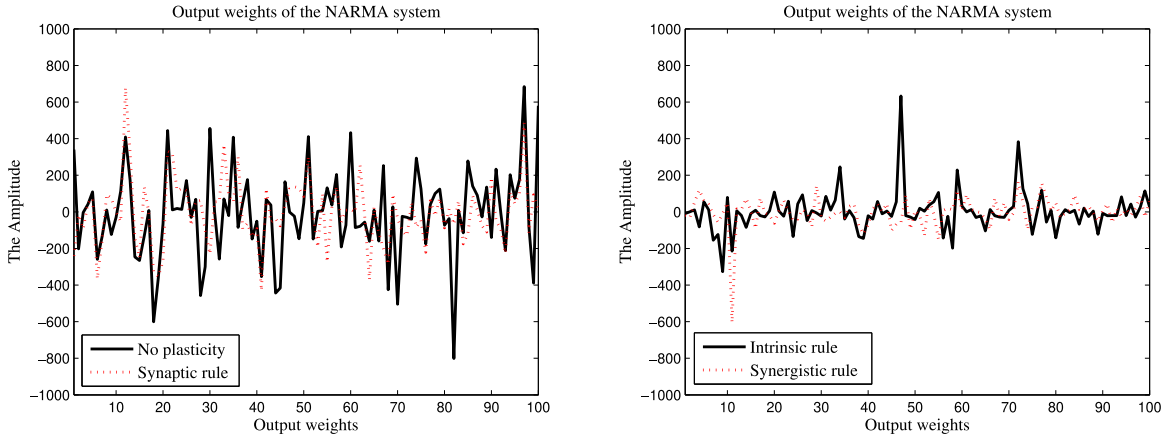

Fig. 6. Output weights using different plasticity rules on the NARMA system.

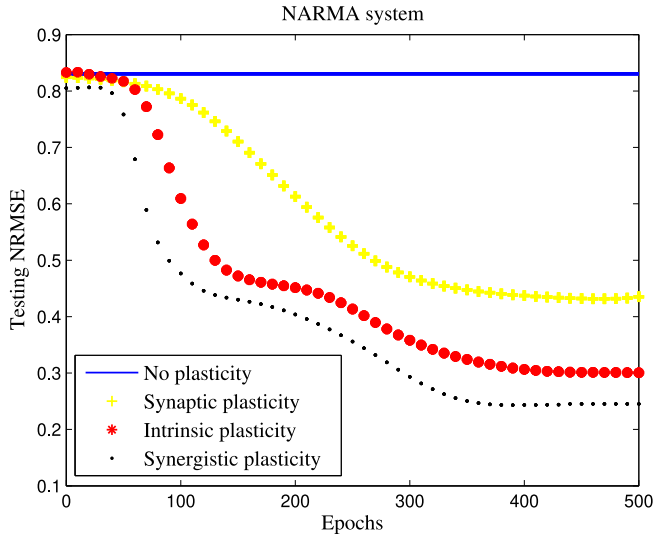

Fig. 7. Convergence curves of the three learning rules on the NARMA system.

properties of the echo state networks, initial connection weights of the reservoir W_0^{res} are usually scaled by $W^{res} = W_0^{res}/\rho$, where ρ is the spectral radius of W_0^{res} [18]. Fig. 3 shows the comparative result for the Lorenz prediction problem by using the original ESN with different spectral radius. The spectral radius is set to 0.8 for the Lorenz prediction problem. In addition, the performance of ESNs with different spectral radius is also compared for other three benchmark problems. For the NARMA system, the Henon map system and the EEG eye state classification task, the spectral radius is all set to 0.8.

4.2.2. The setting of the reservoir size

To set up the reservoir size, the performance of ESNs with different reservoir sizes is compared on each dataset studied in this work. Fig. 4(a) and (b) show the results of the training and testing set for the EEG classification task and the Lorenz prediction

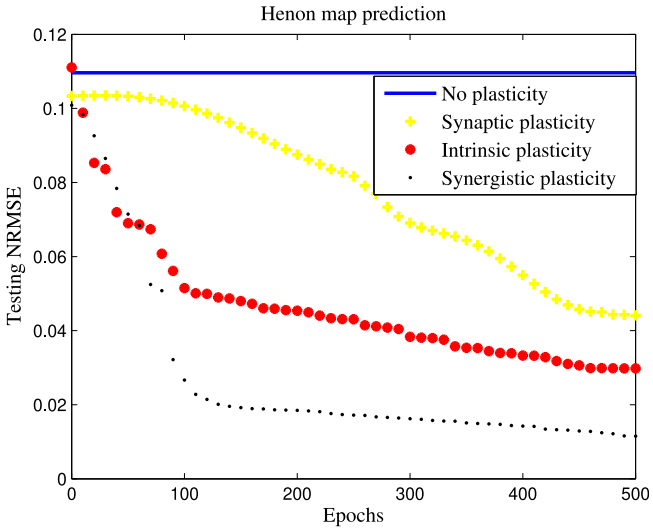

Fig. 8. Learning curves of the four ESN models on the Henon map system.

Table 2
The comparison results of the NARMA system.

Approach	Testing NRMSE (Mean)	Testing NRMSE (Std.)	Number of neurons	Spectral radius
Synergistic rule	0.2463	2.5596e−03	100	0.9324
Synaptic rule	0.4352	3.4150e−03	100	0.9081
Intrinsic rule	0.3004	4.5924e−03	100	0.8000
GESN	0.8289	3.4566e−04	100	0.8000
PESN	0.5957	1.6702e−03	100	0.8000
LESN	0.6377	1.2991e−03	100	0.8000
GRU	0.5530	2.4498e−03	100	0.8000
ESN	0.8306	7.8666e−04	100	0.8000

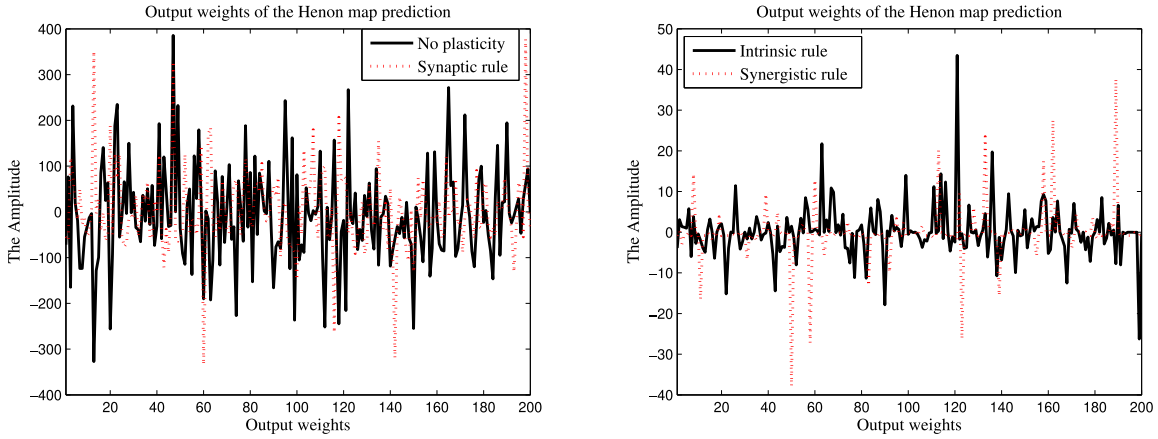

Fig. 9. Output weights using different plasticity rules on the Henon map system.

Table 3
The comparison results of the Henon map system.

Approach	Testing NRMSE (Mean)	Testing NRMSE (Std.)	Number of neurons	Spectral radius
Synergistic rule	0.0115	0.0074	200	0.8225
Synaptic rule	0.0441	0.0137	200	0.8159
Intrinsic rule	0.0298	0.0083	200	0.8000
GESN	0.0787	0.0154	200	0.8000
PESN	0.0745	0.0169	200	0.8000
LESN	0.0353	0.0068	200	0.8000
GRU	0.0385	0.0100	200	0.8000
ESN	0.1096	0.0205	200	0.8000

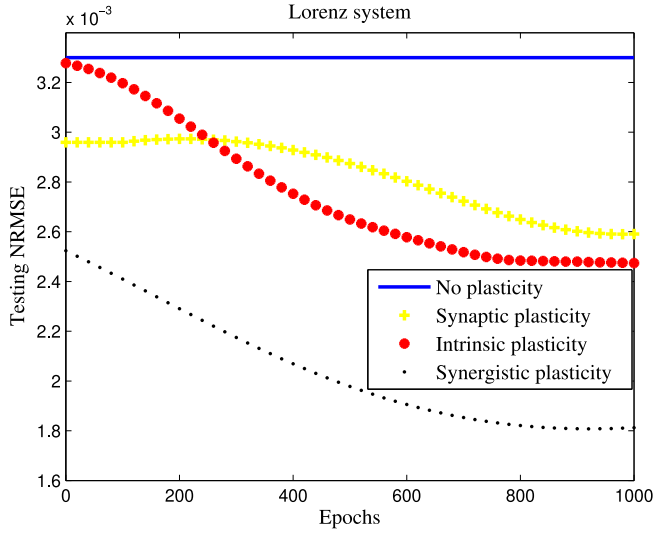

Fig. 10. Convergence curves of the three learning rules on the Lorenz system.

problem by using original ESN with different reservoir sizes, respectively. For the EEG eye state classification task, as shown in Fig. 4(a), the classification accuracy of the training and testing set increases with the increase of the reservoir size. The ESN with 500 reservoir neurons can obtain a good performance, thus the reservoir size is set to 500 for the EEG eye state classification task. For the Lorenz prediction problem, similarly, the reservoir size is set to 200. In addition, the performance of ESNs with different reservoir sizes is also compared on the NARMA and the Henon map system. The reservoir sizes of the NARMA system and the Henon map system are set to 100 and 200, respectively. All experiments have been run independently 30 times for comparison.

4.2.3. The setting of the learning rate

We investigate the effect of the learning rate of neural plasticity rules on the learning performance on the EEG eye state dataset, as shown in Fig. 5. In general, the learning rate should be a small positive constant, as discussed in Sections 2.2 and 2.3. A too large learning rate easily produces an unstable learning process, while a too small learning rate leads to a slow learning speed.

4.3. Algorithms under comparison

To investigate the performance of the proposed plasticity rule, we compare it with the canonical ESN without using any plasticity [5], three state-of-the-art ESN variants, (namely LESN [15], GESN [21] and PESN [16]), and GRU model [68]. For fair comparisons, we adopt the same reservoir sizes and the range of initial weights in all compared algorithms, as discussed in Sections 4.2. In addition, The ESN variants with the synaptic plasticity rule, intrinsic plasticity rule and the synergistic plasticity rule are also compared using the same learning rate for the plasticity rule and the same number of epochs for the same benchmark problem.

5. Experimental results

In the section, we present empirical results of the ESN with synergetic plasticity rule and the compared algorithms on a number of prediction and classification benchmark problems, including the NARMA system, the Henon map system, the Lorenz system and the EEG eye state classification task.

In the following the ESNs with the synaptic plasticity rule, intrinsic plasticity rule and the synergistic plasticity rule are compared. Note that the synaptic plasticity rule can only regulate the connection weights within the reservoir, while the intrinsic plasticity rule can only shape the activation function of an individual neuron. By contrast, the proposed synergistic plasticity rule is able to simultaneously adjust the synaptic weights and the neuronal activation function.

5.1. NARMA system

For the NARMA system, Fig. 7 shows the comparative results of the learning performance of the canonical ESN and its three variants, one with synaptic plasticity, one with intrinsic plasticity, and one with the synergistic plasticity. As Fig. 7 illustrates, both synaptic plasticity and intrinsic plasticity can contribute to improving the learning performance of the ESN compared to the ESN without plasticity. Note that the ESN with the synergistic plasticity rule has obtained the minimum test error (see Table 1).

Fig. 6,7 shows the weights of the output layer of the four ESN variants under comparison. From the left panel of Fig. 6, we can see that the synaptic plasticity rule cannot effectively reduce the amplitude of the output weights compared with the original ESN. However, the intrinsic plasticity rule and synergistic plasticity rule can substantially reduce the amplitude of the output weights, confirming that the intrinsic plasticity can maintain the neuronal firing level in a desirable region and maximize information transmission. In general, small output weights are conductive to improving generalization capabilities of the reservoir [69].

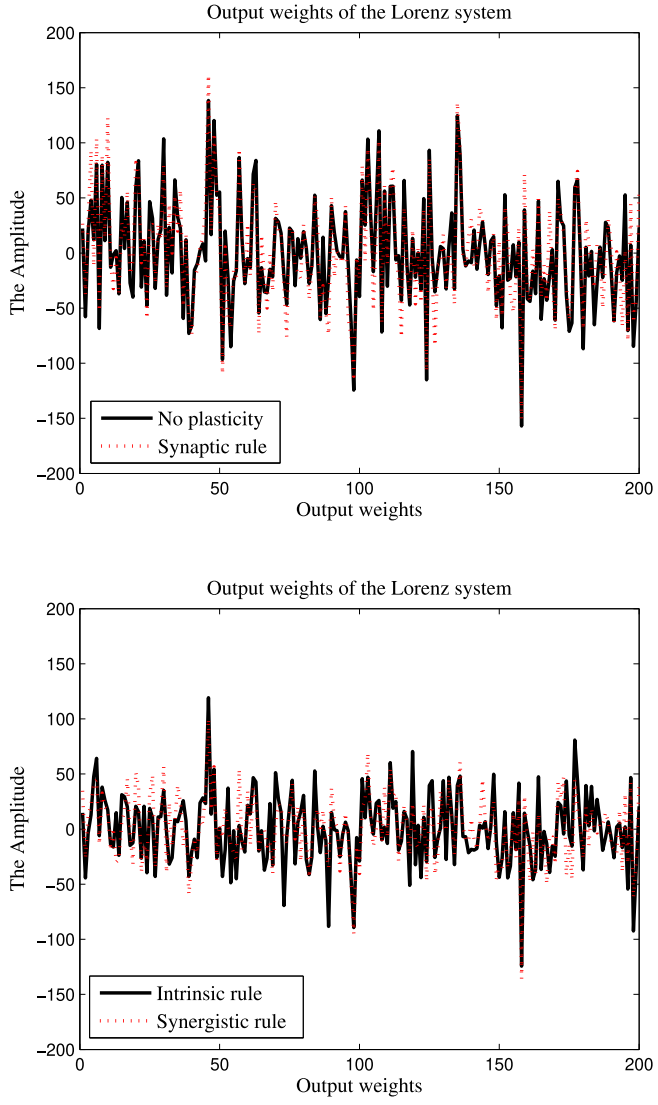

Fig. 11. Output weight of the ESNs using different plasticity rules on the Lorenz system.

Finally, we compare the learning performance of the three ESN variants averaged over 30 independent trials. The test results and network parameters are listed in Table 2. From Table 2 we can see that the ESN using either the synaptic plasticity rule or the intrinsic plasticity rule can improve the learning performance of the ESN. In addition, the connection weights within the reservoir can be regulated by the synaptic plasticity rule, which can also be seen from the results in the change of the spectral radius, as shown in the fifth column of Table 2. However, the intrinsic plasticity can only adapt the neuronal activation function and maintain neuron's homeostasis, which does not affect the spectral radius of the reservoir. For the GRU model, the spectral radius is the mean value of the largest absolute of the eigenvalue of the connection weight of three gate units. In addition, the ESN with the proposed synergistic learning rule also performs better than GESN, PESN, LESN and GRU. The detailed comparative results of the different methods are presented in Table 2.

5.2. Henon map prediction

Similarly, the original ESN and its three variants are compared on the Henon map system. The test results are plotted in Fig. 8. These results clearly show that the proposed synergistic plasticity rule performs better than the existing synaptic plasticity rule and intrinsic plasticity rule.

Additionally, Fig. 9 shows the output weights of the four ESN variants under comparison. From the left panel of Fig. 9, we can see that the synaptic plasticity rule cannot reduce the amplitude of output weights. By contrast, as shown in the right panel of Fig. 9, the ESN variants with either the intrinsic plasticity rule or synergistic plasticity rule can significantly reduce the amplitude of the output weights, which can enhance the generalization capabilities in the presence of different input signals.

The test results and network parameters averaged over 30 independent runs are listed in Table 3. We can see from the table that the ESN with the synergistic plasticity rule has achieved the minimum NRMSE on the test data.

5.3. Lorenz system

For the Lorenz system, similar experimental results can be obtained. As seen from Fig. 10, the ESN with a single neural plasticity rule contribute to enhance the learning performance of the network compared with the original ESN. At the same time, the ESN with the proposed synergistic plasticity rule can further improve the test performance of the ESN.

Fig. 11 shows the resulting output weights of the ESN variants employing three different plasticity learning rules. As shown in Fig. 11, we can also find that the amplitude of output weights of the ESN with the synaptic plasticity rule and the original ESN are roughly the same. In addition, both the intrinsic plasticity rule and the synergistic plasticity rule are able to reduce the output weights.

The test results and the selected parameters of the compared algorithms are summarized in Table 4. As shown in Table 4, the ESN with the synergistic learning rule performs the best on the Lorenz system among seven ESN models and GRU model under comparison.

Table 4
The comparative results on the Lorenz system.

Approach	Testing NRMSE (Mean)	Testing NRMSE (Std.)	Number of neurons	Spectral radius
Synergistic rule	0.0018	5.4846e−04	200	0.8335
Synaptic rule	0.0026	7.1000e−04	200	0.8289
Intrinsic rule	0.0025	6.0802e−04	200	0.8000
GESN	0.0028	4.8602e−04	200	0.8000
PESN	0.0029	1.1629e−03	200	0.8000
LESN	0.0031	6.5181e−04	200	0.8000
GRU	0.0026	4.5785e−04	200	0.8000
ESN	0.0033	6.6531e−04	200	0.8000

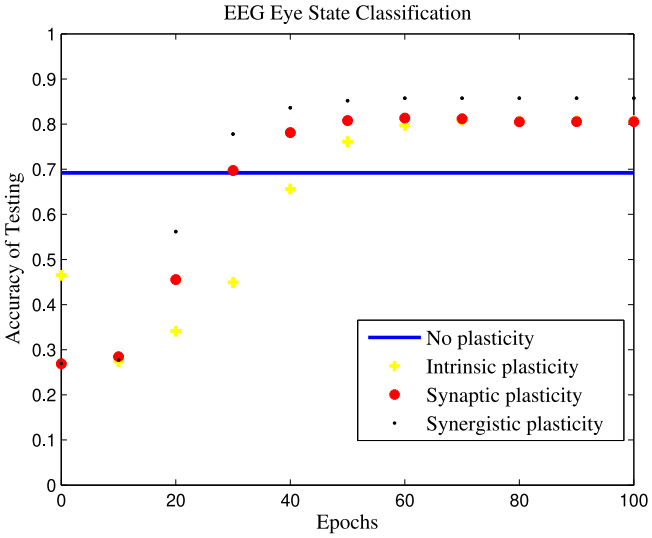

Fig. 12. Convergence curves of the three learning rules on the EEG eye state dataset.

5.4. EEG eye state classification

For the classification task, similar experimental results are obtained, as seen from Fig. 12. At the same time, the ESN with the synergistic plasticity rule can further improve recognition accuracy on the test dataset compared to the ESN with the synaptic plasticity and the intrinsic plasticity, respectively.

The output weights of the compared ESNs on the EEG eye state task are presented in Fig. 13. No clear change to the output weights can be observed from the results of the ESN with synaptic plasticity compared to the original ESN, as shown in the left panel of Fig. 13. By contrast, the ESN with either the intrinsic plasticity rule or the synergistic plasticity rule can reduce the output weights, which can be clearly seen from the right panel of Fig. 13. In other words, both the intrinsic and synergistic plasticity rules can enhance the generalization capabilities by adapting to changes in input.

Finally, the test results and selected parameters of different methods are summarized in Table 5. As shown in Table 5, the ESN with the synergistic learning rule performs the best among the seven compared ESN variants and GRU model.

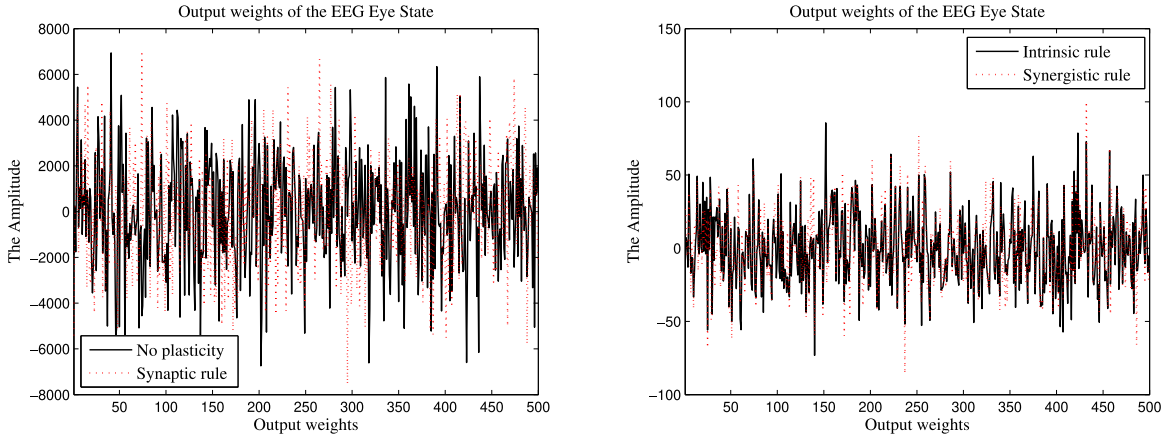

Fig. 13. Output weight by using different plasticity rules on the EEG eye state classification.

Table 5
The comparison results of the EEG eye state classification.

Approach	Accuracy (Mean)	Accuracy (Std.)	Number of neurons	Spectral radius
Synergistic rule	0.8573	0.0201	500	0.9604
Synaptic rule	0.8056	0.0158	500	0.9445
Intrinsic rule	0.8073	0.0285	500	0.8000
GESN	0.7964	0.0241	500	0.8000
PESN	0.7114	0.0199	500	0.8000
LESN	0.7041	0.0205	500	0.8000
GRU	0.7469	0.0181	500	0.8000
ESN	0.6924	0.0160	500	0.8000

6. Conclusion

As two most commonly seen neural plasticity mechanisms, synaptic plasticity and intrinsic plasticity play important roles in increasing the adaptation capability of the randomized learning models and maintaining an individual neuron's homeostasis of its mean output. Although neurobiological studies indicated that changes in both synaptic strength and neuronal intrinsic excitability will directly alter circuit dynamics, no computational models that integrate both neural plasticity rules have been reported. In this paper, a synergistic plasticity learning rule simultaneously considering the regulation of the synaptic weights using the synaptic plasticity rule and the adjustment of neuronal intrinsic excitability using the intrinsic plasticity is proposed. The ESN with the proposed synergistic plasticity rule is compared with the canonical ESN, the ESN with synaptic plasticity, the ESN with intrinsic plasticity, the GRU model, as well as three state-of-the-art ESN extensions on three time series prediction tasks and one classification task. Our experimental results demonstrate that the ESN with the synergistic plasticity learning rule performs the best.

This work focuses on two neural plasticity rules, although biological nervous systems contain multiple forms of neural plasticity mechanisms. A systematic combination of the multiple forms of neural plasticity mechanisms may further improve the learning performance of the ESN. In addition, neuromodulation and some plasticity-related proteins could also modulate the synaptic strength and neuronal intrinsic excitability. Our future work will consider using multiple plasticity mechanisms and neuromodulation for tuning the weights of the reservoir connections. Finally, along the line of research in [7], we will consider combining neural plasticity rules with computationally efficient surrogate-assisted optimization methods [70] to search for optimal architectures of large-scale ESNs for accomplishing more complex tasks.

Declaration of Competing Interest

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

Acknowledgments

This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant CUSF-DH-D-2018101 and Grant 2232016D-32, in part by the National Nature Science Foundation of China under Grant 61473078, Grant 61503075, and Grant 61603090, in part by the International Collaborative Project of the Shanghai Committee of Science and Technology under Grant 16510711100, in part by the Agricultural Project of the Shanghai Committee of Science and Technology under Grant 16391902800, in part by the Shanghai Science and Technology Promotion Project from Shanghai Municipal Agriculture Commission under Grant 2016-1-5-12.

References

[1] A. Rodan, P. Tino, Minimum complexity echo state network, IEEE Trans. Neural Netw. 22 (1) (2010) 131–144.
[2] B. Schrauwen, M. Wardermann, D. Verstraeten, J.J. Steil, D. Stroobandt, Improving reservoirs using intrinsic plasticity, Neurocomputing 71 (7–9) (2008) 1159–1171.
[3] H. Duan, X. Wang, Echo state networks with orthogonal pigeon-inspired optimization for image restoration, IEEE Trans. Neural Netw. Learn. Syst. 27 (11) (2015) 2413–2425.
[4] Z.K. Malik, A. Hussain, Q.J. Wu, Multilayered echo state machine: a novel architecture and algorithm, IEEE Trans. Cybern. 47 (4) (2016) 946–959.
[5] H. Jaeger, The 'echo state' approach to analysing and training recurrent neural networks-with an erratum note, German Nat. Res. Center Inf. Technol., Bonn, Germany, Tech. Rep. GMD 148 (2001)..
[6] W. Maass, T. Natschläger, H. Markram, Real-time computing without stable states: a new framework for neural computation based on perturbations, Neural Comput. 14 (11) (2002) 2531–2560.
[7] Y. Zhou, Y. Jin, J. Ding, Surrogate-assisted evolutionary search of spiking neural architectures in liquid state machines, Neurocomputing 406 (2020) 12–23.
[8] M. Han, M. Xu, Laplacian echo state network for multivariate time series prediction, IEEE Trans. Neural Netw. Learn. Syst. 29 (1) (2017) 238–244.
[9] Q. Ma, L. Shen, W. Chen, J. Wang, J. Wei, Z. Yu, Functional echo state network for time series classification, Inf. Sci. 373 (2016) 1–20.
[10] L. Sun, B. Jin, H. Yang, J. Tong, C. Liu, H. Xiong, Unsupervised EEG feature extraction based on echo state network, Inf. Sci. 475 (2019) 1–17.
[11] L. Bozhkov, P. Koprinkova-Hristova, P. Georgieva, Reservoir computing for emotion valence discrimination from EEG signals, Neurocomputing 231 (2017) 28–40.
[12] J. Yin, Y. Meng, Y. Jin, A developmental approach to structural self-organization in reservoir computing, IEEE Trans. Auton. Mental Dev. 4 (4) (2012) 273–289.
[13] H. Jaeger, H. Haas, Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication, Science 304 (5667) (2004) 78–80.
[14] Q. Wu, E. Fokoue, D. Kudithipudi, On the statistical challenges of echo state networks and some potential remedies, arXiv preprint arXiv:1802.07369..
[15] H. Jaeger, M. Lukoševičius, D. Popovici, U. Siewert, Optimization and applications of echo state networks with leaky-integrator neurons, Neural Netw. 20 (3) (2007) 335–352.
[16] C. Yang, J. Qiao, H. Han, L. Wang, Design of polynomial echo state networks for time series prediction, Neurocomputing 290 (2018) 148–160.
[17] F.M. Bianchi, L. Livi, C. Alippi, Investigating echo-state networks dynamics by means of recurrence analysis, IEEE Trans. Neural Netw. Learn. Syst. 29 (2) (2016) 427–439.
[18] Y. Kawai, J. Park, M. Asada, A small-world topology enhances the echo state property and signal propagation in reservoir computing, Neural Netw. 112 (2019) 15–23.
[19] D. Kudithipudi, Q. Saleh, C. Merkel, J. Thesing, B. Wysocki, Design and analysis of a neuromemristive reservoir computing architecture for biosignal processing, Front. Neurosci. 9 (2016) 502.
[20] Q. Ma, E. Chen, Z. Lin, J. Yan, Z. Yu, W.W. Ng, Convolutional multitimescale echo state network, IEEE Trans. Cybern..
[21] J. Qiao, F. Li, H. Han, W. Li, Growing echo-state network with multiple subreservoirs, IEEE Trans. Neural Netw. Learn. Syst. 28 (2) (2016) 391–404.
[22] M. Xu, Y. Yang, M. Han, T. Qiu, H. Lin, Spatio-temporal interpolated echo state network for meteorological series prediction, IEEE Trans. Neural Netw. Learn. Syst. 30 (6) (2019) 1621–1634.
[23] L.A. Thiede, U. Parlitz, Gradient based hyperparameter optimization in echo state networks, Neural Netw. 115 (2019) 23–29.
[24] J. Yperman, T. Becker, Bayesian optimization of hyper-parameters in reservoir computing, arXiv preprint arXiv:1611.05193..
[25] K.C. Chatzidimitriou, P.A. Mitkas, A NEAT way for evolving echo state networks, ECAI (2010) 909–914.
[26] N. Chouikhi, R. Fdhila, B. Ammar, N. Rokbani, A.M. Alimi, Single-and multi-objective particle swarm optimization of reservoir structure in echo state network, in: 2016 International Joint Conference on Neural Networks (IJCNN), IEEE, 2016, pp. 440–447.
[27] M.-H. Yusoff, J. Chrol-Cannon, Y. Jin, Modeling neural plasticity in echo state networks for classification and regression, Inf. Sci. 364 (2016) 184–196.
[28] A. Destexhe, E. Marder, Plasticity in single neuron and circuit computations, Nature 431 (7010) (2004) 789–795.
[29] J. Chrol-Cannon, Y. Jin, Computational modeling of neural plasticity for self-organization of neural networks, BioSystems 125 (2014) 43–54.
[30] R. Mozzachiodi, J.H. Byrne, More than synaptic plasticity: role of nonsynaptic plasticity in learning and memory, Trends Neurosci. 33 (1) (2010) 17–26.
[31] A. Citri, R.C. Malenka, Synaptic plasticity: multiple forms, functions, and mechanisms, Neuropsychopharmacology 33 (1) (2008) 18–41.
[32] D.O. Hebb, The Organization of Behavior: A Neuropsychological Theory, Wiley, New York, USA, 1949.
[33] E.L. Bienenstock, L.N. Cooper, P.W. Munro, Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex, J. Neurosci. 2 (1) (1982) 32–48.
[34] Š. Babinec, J. Pospíchal, Improving the prediction accuracy of echo state neural networks by anti-Oja's learning, in: International Conference on Artificial Neural Networks, Springer, 2007, pp. 19–28..
[35] G.-Q. Bi, M.-M. Poo, Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type, J. Neurosci. 18 (24) (1998) 10464–10472.
[36] W. Zhang, D.J. Linden, The other side of the engram: experience-driven changes in neuronal intrinsic excitability, Nat. Rev. Neurosci. 4 (11) (2003) 885–900.
[37] J. Triesch, A gradient rule for the plasticity of a neuron's intrinsic excitability, in: International Conference on Artificial Neural Networks, Springer, 2005, pp. 65–70.
[38] C. Li, Y. Li, A spike-based model of neuronal intrinsic plasticity, IEEE Trans. Auton. Mental Dev. 5 (1) (2012) 62–73.
[39] P. Koprinkova-Hristova, G. Palm, ESN intrinsic plasticity versus reservoir stability, in: International Conference on Artificial Neural Networks, Springer, 2011, pp. 69–76..
[40] P. Koprinkova-Hristova, N. Tontchev, Echo state networks for multi-dimensional data clustering, in: International Conference on Artificial Neural Networks, Springer, 2012, pp. 571–578.

[41] P. Koprinkova-Hristova, On effects of IP improvement of ESN reservoirs for reflecting of data structure, in: 2015 International Joint Conference on Neural Networks (IJCNN), IEEE, 2015, pp. 1–7..
[42] R. Baddeley, L.F. Abbott, M.C. Booth, F. Sengpiel, T. Freeman, E.A. Wakeman, E.T. Rolls, Responses of neurons in primary and inferior temporal visual cortices to natural scenes, Proc. Roy. Soc. Lond. Ser. B Biol. Sci. 264 (1389) (1997) 1775–1783.
[43] G. Scheler, Learning intrinsic excitability in medium spiny neurons, F1000Res. 2..
[44] X. Wang, Y. Jin, K. Hao, Echo state networks regulated by local intrinsic plasticity rules for regression, Neurocomputing 351 (2019) 111–122.
[45] T.C. Francis, S.C. Gantz, K. Moussawi, A. Bonci, Synaptic and intrinsic plasticity in the ventral tegmental area after chronic cocaine, Curr. Opin. Neurobiol. 54 (2019) 66–72.
[46] M.E. Lambo, G.G. Turrigiano, Synaptic and intrinsic homeostatic mechanisms cooperate to increase L2/3 pyramidal neuron excitability during a late phase of critical period plasticity, J. Neurosci. 33 (20) (2013) 8810–8819.
[47] A.H. Gittis, S. du Lac, Intrinsic and synaptic plasticity in the vestibular system, Curr. Opin. Neurobiol. 16 (4) (2006) 385–390.
[48] C.L. Darlington, M.B. Dutia, P.F. Smith, The contribution of the intrinsic excitability of vestibular nucleus neurons to recovery from vestibular damage, Eur. J. Neurosci. 15 (11) (2002) 1719–1727.
[49] G. Scheler, Neuromodulation influences synchronization and intrinsic read-out, F1000Res. 7..
[50] J. Triesch, Synergies between intrinsic and synaptic plasticity in individual model neurons, in: Advances in Neural Information Processing Systems, 2005, pp. 1417–1424..
[51] Y. Li, C. Li, Synergies between intrinsic and synaptic plasticity based on information theoretic learning, PloS One 8 (5) (2013) e62894.
[52] M. Lukoševičius, H. Jaeger, Reservoir computing approaches to recurrent neural network training, Comput. Sci. Rev. 3 (3) (2009) 127–149.
[53] H. Jaeger, Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the echo state network approach, Ger. Nat. Res. Center Inf. Technol. Sankt August. Ger. Tech. Rep., 2002.
[54] E. Benito, A. Barco, CREB's control of intrinsic and synaptic plasticity: implications for CREB-dependent memory models, Trends Neurosci. 33 (5) (2010) 230–240.
[55] J. Tan, C. Quek, A BCM theory of meta-plasticity for online self-reorganizing fuzzy-associative learning, IEEE Trans. Neural Netw. 21 (6) (2010) 985–1003.
[56] A. Carlson, Anti-Hebbian learning in a non-linear neural network, Biol. Cybern. 64 (2) (1990) 171–176.
[57] X. Wang, Y. Jin, K. Hao, Evolving local plasticity rules for synergistic learning in echo state networks, IEEE Trans. Neural Netw. Learn. Syst. 31 (4) (2019) 1363–1374.
[58] E. Oja, Simplified neuron model as a principal component analyzer, J. Math. Biol. 15 (3) (1982) 267–273.
[59] J.D. Shepherd, G. Rumbaugh, J. Wu, S. Chowdhury, N. Plath, D. Kuhl, R.L. Huganir, P.F. Worley, Arc/arg3. 1 mediates homeostatic synaptic scaling of ampa receptors, Neuron 52 (3) (2006) 475–484.
[60] W.C. Abraham, M.F. Bear, Metaplasticity: the plasticity of synaptic plasticity, Trends Neurosci. 19 (4) (1996) 126–130.
[61] A.F. Atiya, A.G. Parlos, New results on recurrent network training: unifying the algorithms and accelerating convergence, IEEE Trans. Neural Netw. 11 (3) (2000) 697–709.
[62] M. Hénon, A two-dimensional mapping with a strange attractor, in: The Theory of Chaotic Attractors, Springer, 1976, pp. 94–102.
[63] P. Ping, F. Xu, Y. Mao, Z. Wang, Designing permutation–substitution image encryption networks with Henon map, Neurocomputing 283 (2018) 53–63.
[64] J. Lu, R. Wei, X. Wang, Z. Wang, Backstepping control of discrete-time chaotic systems with application to the henon system, IEEE Trans. Circ. Syst. 48 (11) (2001) 1359–1363.
[65] W. Song, J. Liang, Difference equation of lorenz system, Int. J. Pure Appl. Math. 83 (1) (2013) 101–110.
[66] A. Frank, UCI machine learning repository, http://archive.ics.uci.edu/ml..
[67] T. Wang, S.-U. Guan, K.L. Man, T. Ting, Eeg eye state identification using incremental attribute learning with time-series classification, Math. Probl. Eng. (2014).
[68] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using rnn encoder-decoder for statistical machine translation, in: Proceeding of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25, 2014 – October 29, 2014, Doha, Qatar, Association for Computational Linguistics (ACL), 2014, pp. 1724–1734..
[69] H. Jaeger, Reservoir riddles: suggestions for echo state network research, in: Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005, vol. 3, IEEE, 2005, pp. 1460–1462..
[70] C. Sun, Y. Jin, R. Cheng, J. Ding, J. Zeng, Surrogate-assisted cooperative swarm optimization of high-dimensional expensive problems, IEEE Trans. Evol. Comput. 21 (4) (2017) 644–660.

Xinjie Wang was born in Suzhou, Anhui Province, China, in 1991. He obtained his Ph. D. degree with the College of Information Sciences and Technology, Donghua University, Shanghai, China in 2020 and is currently a postdoctoral research associate in the Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, China. His current research interests include computational modeling of plasticity and data driven evolutionary optimization.

Yaochu Jin received the B.Sc., M.Sc., and Ph.D. degrees from Zhejiang University, Hangzhou, China, in 1988, 1991, and 1996 respectively, and the Dr.-Ing. degree from Ruhr University Bochum, Germany, in 2001.
He is a Distinguished Chair Professor in Computational Intelligence, Department of Computer Science, University of Surrey, Guildford, U.K. He was also a Finland Distinguished Professor, University of Jyvaskyla, Finland and a Changjiang Distinguished Professor, Northeastern University, China. His main research interests include evolutionary computation, machine learning, computational neuroscience, and evolutionary developmental systems, with their application to data-driven optimization and decision-making, self-organizing swarm robotic systems, and bioinformatics. He has (co)authored over 300 peer-reviewed journal and conference papers and has been granted eight patents on evolutionary optimization.
Dr Jin is the Editor-in-Chief of the IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS and Complex & Intelligent Systems. He was an IEEE Distinguished Lecturer (2013-2015, 207-2019) and was Vice President for Technical Activities of the IEEE Computational Intelligence Society. He is a Fellow of IEEE.

Kuangrong Hao (M?17) received the B.S. and M.S. degrees in mechanical engineering from Hebei University of Technology, Tianjin, China in 1984 and 1989, respectively. She received the M. S. and Ph.D. degrees in applied mathematics and computer sciences from Ecole Normale Sup?rieur de Cachan, France in 1991, and Ecole Nationale des Ponts et Chauss?es, Paris, France in 1995, respectively.
She is currently a Professor at the College of Information Sciences and Technology, Donghua University, Shanghai, China. She has published more than 200 technical papers, more than 50 granted invention patents, and five research monographs. Her scientific interests include intelligent perception, intelligent systems and network intelligence, robot control, and intelligent optimization of textile industrial process.