

Exploring reservoir computing: Implementation via double stochastic nanowire networks

Jian-Feng Tang(唐健峰), Lei Xia(夏磊), Guang-Li Li(李广隶), Jun Fu(付军), Shukai Duan(段书凯), and Lidan Wang(王丽丹)
Citation: Chin. Phys. B, 2024, 33 (**3**): 037302. DOI: 10.1088/1674-1056/aceeea
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Electrically controllable spin filtering in zigzag phosphorene nanoribbon based normal—antiferromagnet—normal junctions

Ruigang Li(李锐岗), Jun-Feng Liu(刘军丰), and Jun Wang(汪军)
Chin. Phys. B, 2024, 33 (**1**): 017304. DOI: 10.1088/1674-1056/acf82a

Resistive switching properties of SnO_2 nanowires fabricated by chemical vapor deposition

Ya-Qi Chen(陈亚琦), Zheng-Hua Tang(唐政华), Chun-Zhi Jiang(蒋纯志), and De-Gao Xu(徐德高)
Chin. Phys. B, 2023, 32 (**9**): 097302. DOI: 10.1088/1674-1056/acc3fd

Nonlinear current response and electric quantum oscillations in the Dirac semimetal Cd_3As_2

Hao-Nan Cui(崔浩楠), Ze-Nan Wu(吴泽南), Jian-Kun Wang(王建坤), Guang-Yu Zhu(祝光宇), Jia-Jie Yang(杨佳洁), Wen-Zhuang Zheng(郑文壮), Zhi-Min Liao(廖志敏), Shuo Wang(王硕), Ben-Chuan Lin(林本川), and Dapeng Yu(俞大鹏)
Chin. Phys. B, 2023, 32 (**8**): 087306. DOI: 10.1088/1674-1056/acd9c5

Negative magnetoresistance in Dirac semimetal Cd_3As_2 with in-plane magnetic field perpendicular to current

Hao-Nan Cui(崔浩楠), Guang-Yu Zhu(祝光宇), Jian-Kun Wang(王建坤), Jia-Jie Yang(杨佳洁), Wen-Zhuang Zheng(郑文壮), Ben-Chuan Lin(林本川), Zhi-Min Liao(廖志敏), Shuo Wang(王硕), and Da-Peng Yu(俞大鹏)
Chin. Phys. B, 2023, 32 (**7**): 077305. DOI: 10.1088/1674-1056/accd50

Conductive path and local oxygen-vacancy dynamics: Case study of crosshatched oxides

Z W Liang(梁正伟), P Wu(吴平), L C Wang(王利晨), B G Shen(沈保根), and Zhi-Hong Wang(王志宏)
Chin. Phys. B, 2023, 32 (**4**): 047303. DOI: 10.1088/1674-1056/acb421

Exploring reservoir computing: Implementation via double stochastic nanowire networks

Jian-Feng Tang(唐健峰)[1,6], Lei Xia(夏磊)[1], Guang-Li Li(李广隶)[1], Jun Fu(付军)[1], Shukai Duan(段书凯)[1,3,4], and Lidan Wang(王丽丹)[1,2,3,5,6,†]

[1] *College of Artificial Intelligence, Southwest University, Chongqing* 400715, *China*
[2] *Brain-inspired Computing & Intelligent Control of Chongqing Key Laboratory, Chongqing* 400715, *China*
[3] *National & Local Joint Engineering Laboratory of Intelligent Transmission and Control Technology, Chongqing* 400715, *China*
[4] *Chongqing Brain Science Collaborative Innovation Center, Chongqing* 400715, *China*
[5] *Key Laboratory of Luminescence Analysis and Molecular Sensing, Ministry of Education, Southwest University, Chongqing* 400715, *China*
[6] *State Key Laboratory of Intelligent Vehicle Safety Technology, Chongqing* 400715, *China*

(Received 3 July 2023; revised manuscript received 6 August 2023; accepted manuscript online 10 August 2023)

Neuromorphic computing, inspired by the human brain, uses memristor devices for complex tasks. Recent studies show that self-organizing random nanowires can implement neuromorphic information processing, enabling data analysis. This paper presents a model based on these nanowire networks, with an improved conductance variation profile. We suggest using these networks for temporal information processing via a reservoir computing scheme and propose an efficient data encoding method using voltage pulses. The nanowire network layer generates dynamic behaviors for pulse voltages, allowing time series prediction analysis. Our experiment uses a double stochastic nanowire network architecture for processing multiple input signals, outperforming traditional reservoir computing in terms of fewer nodes, enriched dynamics and improved prediction accuracy. Experimental results confirm the high accuracy of this architecture on multiple real-time series datasets, making neuromorphic nanowire networks promising for physical implementation of reservoir computing.

Keywords: double-layer stochastic (DS) nanowire network architecture, neuromorphic computation, nanowire network, reservoir computing, time series prediction

PACS: 73.63.–b, 05.45.–a **DOI:** 10.1088/1674-1056/aceeea

1. Introduction

In recent years, neuromorphic computing has emerged as a promising technology in the post-Moore's law era, attracting significant attention due to its high degree of connectivity, parallelism and relatively low power consumption.[1] The primary objective of neuromorphic computing is to overcome the main limitation of traditional von Neumann computing architectures; specifically, the data transfer bottleneck known as the von Neumann bottleneck. The progress of neuromorphic computing and engineering depends on the development of new hardware technologies and architectures that mimic the functions of biological neuron circuits. Memristor devices, which can store and process data in the same physical location, play a vital role in bio-inspired and neuromorphic computing. Devices with memristor characteristics can imitate biological synaptic functions, emulate efficient human brain functions and provide hardware support for realizing trusted architectures. Numerous studies have reported the use of devices with memristor properties to achieve artificial synapses, including resistive random-access memory with drifting[2] and diffusive[3] memristors. The paper[4] reports on the implementation of a two-dimensional memristor-based sensor using tin sulfide (SnS) for language learning. It demonstrated an accuracy of 91% in classifying short sentences, showcasing a low-cost, real-time solution suitable for time and sequence signal processing in edge machine learning applications. In the study,[5] a memristor-based reservoir computing system is presented, utilizing the internal short-term ion dynamics of memristive devices and the concept of virtual nodes. It successfully demonstrates time series analysis and prediction tasks. In the research,[6] ultralow-power, high-precision image processing and dynamic video analysis are achieved by utilizing a nonlinear, spatiotemporal correlated photocurrent to construct an internal reservoir computing system within a sensor. Additionally, many studies have utilized memristor devices as modules to simulate synaptic functions[7,8] and research has shown that forming memristor devices in large-scale crossbar arrays for the realization of artificial neural networks and neuromorphic computing is one of the most promising approaches at present.[8]

However, the hardware architecture of regular arrays of amnestic devices is quite different from the intrinsic structural complexity of the human brain and cannot simulate the topology and emergent behavior of biological neural networks. Consequently, researchers have focused on another type of hardware neural network based on memristor synaptic devices, with nanowire networks emerging as a unique representative class. Nanowires network consists of randomly arranged

[†]Corresponding author. E-mail: ldwang@swu.edu.cn
© 2024 Chinese Physical Society and IOP Publishing Ltd

wires, no more than 10 μm long and no thicker than 500 nm, in a two-dimensional plane. Where the wires overlap in the network, they form an electrochemical node resembling a synapse between neurons. Hochstetter *et al.* has proposed a neuro morphodynamic model for nanowire networks that adapt in a brain-like manner.[9] The electrical signals passing through this network automatically find the best path to transmit information and remember previous paths through the system.[10] Professor Tomonobu Nakayama has extensively researched nanowire networks, including their dynamic behavior and network topology. He has demonstrated that nanowire networks respond adaptively to time-varying stimuli, exhibit diverse dynamics that can be tuned from ordered to chaotic and illustrated that 'avalanche criticality' and 'chaotic edge criticality' in nanowire networks can optimize increasingly complex learning tasks.[11] The nanowire network employs a top–down nanofabrication approach, allowing for higher interconnectivity of amnestic crosspoint junctions. In terms of structure and functional relationships, the nanowire network has strong similarities to real (bio)neural networks formed by bottom–up self-assembly. Meanwhile, self-organizing random nanowire networks composed of nano-objects have recently been demonstrated for the physical implementation of neuromorphic information processing.[9,10,12,13] This random self-organized structure is more consistent with the biologically based neuromorphic architecture than the regular memristor array. Fu *et al.*[14,15] demonstrated the ability of a self-organized random nanowire network to implement reservoir computation.

In reservoir computing models, the internal reservoir units are determined by a random sparse matrix of connections.[16] The reservoir network maps the input to a high-dimensional dynamic space, with the values of multiple reservoir nodes determining the mapping result. The mapped data is then fed into the readout layer for classification or prediction. Only the readout layer weights need to be trained, significantly reducing computational effort and improving the network's computational efficiency. A self-organizing random nanonetwork is a highly interconnected network of random nanowires that intersect to form nodes with amnestic properties. The network's overall dynamic behavior arises from the collective behavior of multiple nano nodes. Self-organized stochastic nanonetworks have also been shown to exhibit rich neuromorphic dynamics,[17–20] such as short- and long-term plasticity, heterosynaptic plasticity and chaotic edge properties. The structural and functional similarities between self-organized stochastic nanonetworks and reservoir networks justify using self-organized stochastic nanonetworks as an internal implementation of reservoir computing.

Hochstetter *et al.*[11] and Zhu *et al.*[9] introduced a random nanowire network model, suggesting that the network's size and complexity are associated with the average variance and length of the nanowires. The model employs a threshold equation to represent the dynamics of each memristor edge and assumes uniform behavior across all nanowire junctions in the network. Ricciardi *et al.*[21] presented a grid graphical modelling method for memristor nanowire networks, where a physics-based accretion–inhibition rate balance equation governs each edge's dynamic behavior. However, due to the randomness of mechanical contact, the resistive state and memristor effect of nanowire nodes are randomly distributed, rendering the model incapable of accurately representing the network's dynamic behavior.

In this paper, we design a nanowire network model that better aligns with the hardware characteristics of the process. By combining the features of nanowire networks, we propose converting data into voltage pulses of varying amplitudes as input to the network, thereby reducing information processing time and enhancing computational efficiency. We develop a double-layer stochastic (DS) reservoir network based on the nanowire network, utilizing the random distribution of voltage states at the nodes and the rich dynamic behavior of the web to construct a reservoir network. This design also incorporates the nanowire network's ability to process multiple input signals in a spatiotemporal manner. We test the DS reservoir network's performance on real-time series data, and our experiments demonstrate the superior accuracy and stability of the nanowire network-based reservoir system compared to traditional reservoir networks. In terms of energy consumption, a study conducted by Milano *et al.* in 2022 found that the total energy consumption of a single input is 750 μJ, which is higher than the energy consumption of a single memory resistor device in traditional top–down reservoir computing systems. However, by utilizing shorter pulse widths and adjusting material properties, it is possible to further reduce energy consumption, thus providing potential for nanowire networks in future complementary metal–oxide–semiconductor integration.[13]

2. Network modeling and analysis

2.1. Dynamic modeling of memristive resistance for random network graphs

Zegarac *et al.*[22] established a link between graph theory and the dynamics of memristive circuits. Milano *et al.*[23] employed a graph-theoretic approach to investigate the functionality and connectivity of random nanowire networks, introducing the concept of memristive graphs. These graphs represent random nanowire networks as weighted directed graphs, with nodes symbolizing nanowire junctions and connected edges denoting the relationships between them. Montano *et al.*[21] utilized a grid model to simulate nanowire network connections, where nodes are regularly connected, simplifying calcu-

lations and network complexity to some extent. This model offers a general platform for exploring non-traditional computational paradigms in nanowire networks.

However, randomly connected nanowire networks more closely resemble the inter-synaptic connections of brain neurons and the connections within reservoir computing. Therefore, in this paper, we opt for random nanowire network simulation to implement reservoir computing and adopt a graph theory approach to simulate the dynamic behavior of the random nanowire network. Milano *et al.*[17] observed that when a random nanowire network is dynamically stimulated by an external voltage, a process occurs on two outer-wrapped insulating polyvinylpyrrolidone silver nanowires, forming a silver conductive bridge. The formation and rupture of this bridge result in a resistance switch.

The network dynamics rely on the collective behavior of numerous nanowire nodes within the network and can be modelled based on the memristor behavior of a single nanowire node. To simulate the model of short-term plasticity, a single nanowire node is developed, and the network memristor edge dynamics are modelled using a voltage-controlled potential-inhabitance balance equation. The voltage difference between various nanowire nodes in the network and the relationship between the currents flowing between the nodes under this voltage difference is represented by

$$I(t) = \left[G_{\text{low}}(1-g(t)) + G_{\text{high}}(g(t))\right] V(t), \tag{1}$$

where G_{low} and G_{high} denote the minimum and maximum conductance of the memristor edge, respectively; while $g(t)$ represents the normalized conductance of the memristor edge, normalized between 0 and 1. The current relationship of $I(t)$ follows Ohm's first law. The memristor edge conductance variation is expressed by a voltage-controlled potential-inhibit rate balance equation

$$\frac{\mathrm{d}g}{\mathrm{d}t} = \omega_{\text{a}}(1-g) + \omega_{\text{b}} g. \tag{2}$$

In this equation, ω_{a} and ω_{b} are functions of the input voltage controlled by the input voltage, representing the inhibition and enhancement coefficients that control the conductance change, respectively,

$$\omega_{\text{a}} = \omega_{\text{a0}}\, \mathrm{e}^{\eta_{\text{a}} V}, \tag{3}$$

$$\omega_{\text{b}} = \omega_{\text{b0}}\, \mathrm{e}^{-\eta_{\text{b}} V}. \tag{4}$$

Equations (3) and (4) are derived from the physical description of ion diffusion.[24–26]

$$g(t) = g_{t-1}\, \mathrm{e}^{\left[\frac{\omega_{\text{a}}}{\omega_{\text{b}}} - (\omega_{\text{a}}+\omega_{\text{b}})\Delta t\right]} + \frac{\omega_{\text{a}}}{\omega_{\text{a}}+\omega_{\text{b}}}\left[1 - \mathrm{e}^{-(\omega_{\text{a}}+\omega_{\text{b}})\Delta t}\right]. \tag{5}$$

Equation (5) can be solved recursively. Following Miranda *et al.*'s approach[27] for handling continuous functions in a discrete-time domain, our model employs the same concept, allowing the model to represent the network's response to arbitrary input voltages. It is important to note that the initial conductance of the memristor edge is uniform and identical to the initial case conductance. Equation (5) is divided into two parts: the first part signifies the network's historical memory state, while the second part represents the network's constantly changing process under external input voltage stimulation.

2.1.1. Analysis of the dynamic behavior of random nanowire networks

In this study, we explore the changes in conductance within a nanowire network when exposed to constant external voltage stimulation. The network features two external terminals: an input terminal for voltage and an output terminal where voltage remains at 0. As external voltage is applied to the input terminal, a spontaneous connectivity path forms between the input and output terminals, resulting in conductance changes within the network. Figure 1(a) illustrates the time-dependent trajectory of effective conductance changes for a realistic Ag nanowire network with a two-terminal configuration under constant voltage pulse stimulation.[23] Meanwhile, Fig. 1(b) presents the trajectory for a previously designed nanowire model subjected to the same stimulation. When compared to earlier research,[23] the conductance change curve in Fig. 1(b) more closely resembles the actual double-ended configuration of the nanowire network under identical experimental conditions.

The nanowire network's effective conductance is enhanced by constant voltage pulse stimulation, and without external voltage stimulation, the network conductance gradually decreases and spontaneously relaxes. Conductance changes at each edge of the network are governed by the enhancement–inhibition rate balance equation. A detailed comparison of Figs. 1(a) and 1(b) indicates that the experimental results align well with the model results during both the enhancement and relaxation phases. This observation confirms the proposed model's ability to accurately simulate the behavior of real nanowire network dynamics in response to external stimulation.

Furthermore, Fig. 2 illustrates the dynamic emergent behavior of the network under external constant pulse stimulation. The nanowire network's conductance continuously increases in response to external voltage stimulation, progressively forming a conduction path between the voltage input and output terminals. As the external voltage stimulation persists, conduction paths are continuously redistributed, leading to the formation of an increasing number of paths within the network. This redistribution of conduction paths highlights the self-organizing characteristics of the nanowire network.

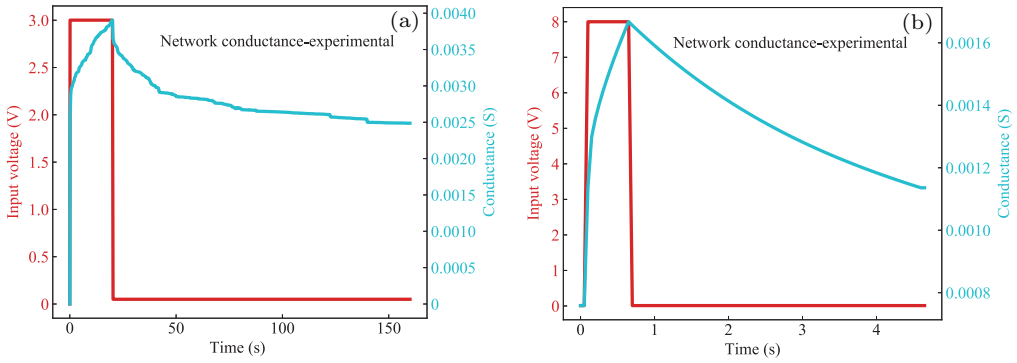


Fig. 1. (a) Plot showing the trajectory of the effective conductance of the Ag nanowire network with time in the experiment under constant voltage pulse stimulation, and (b) plot showing the trajectory of the effective conductance of the network simulated according to the model.

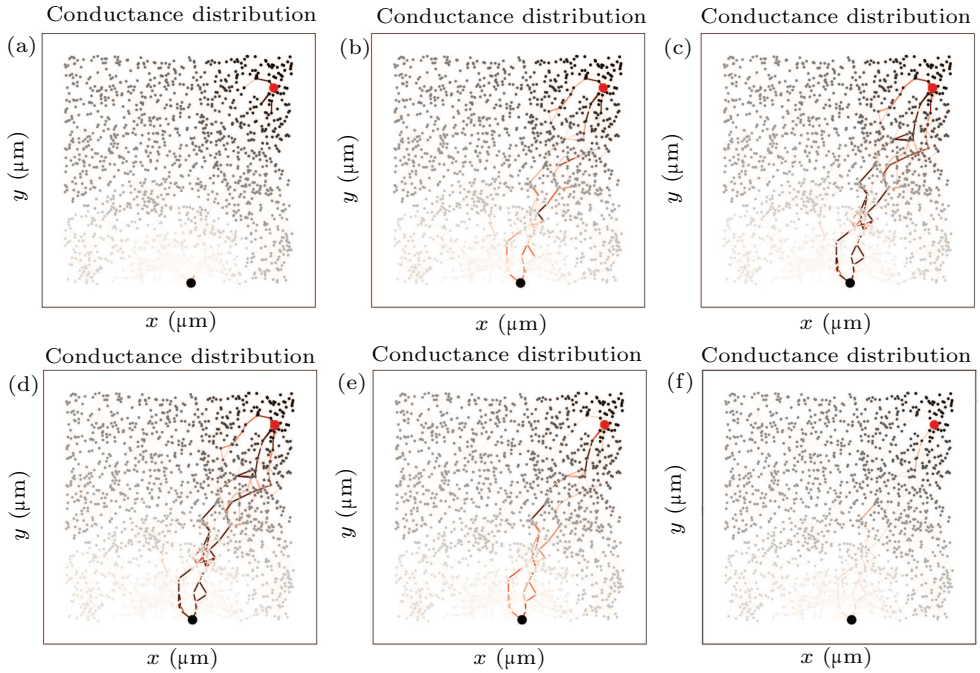


Fig. 2. We can clearly observe the formation and disconnection of conductive paths in the network by examining the evolution of network edge conductance distribution over time under a constant external voltage stimulus. The color shades of the network edges indicate the network edge conductance and the color intensity of the network nodes indicates the node voltage. The network has a voltage input terminal and an output terminal, and the output terminal is grounded.

In the absence of voltage stimulation, the network memristor edge conductance gradually decreases, causing the conduction paths to dissolve and the network conductance to revert to its initial state. The interplay between the enhancement of network conductance under external voltage stimulation and the spontaneous relaxation of the network conductance state enables the network to accumulate information, providing it with the capacity to process complex and data-rich inputs. In a study,[20] the network achieved continuous and adaptive reconfiguration by enhancing the main current pathways and detecting the presence of multiple conduction channels, demonstrating remarkable fault-tolerant capabilities. These silver nanowires self-assembled into a neural-morphic system composed of polymer-coated silver nanowires with resistive switching junctions.

Further research[11] demonstrated that the neural-morphic nanowire network exhibited adaptive and collective responses to electrical stimulation, achieved through a phase transition of threshold voltage and conductivity. Intersection points in the network formed avalanche collective switching, consistent with avalanche phenomena at critical voltage. At high voltages, both large-scale and small-scale avalanches coexisted. Under alternating polarity stimulation, the network could regulate between ordered and chaotic states. These findings suggest that the neural-morphic nanowire network can be tuned to exhibit collective dynamic mechanisms similar to the brain, optimizing information processing. This further underscores the powerful capabilities of nanowire networks in information processing. Moreover, this model simulates the collective dynamic behavior of nanowire networks with reduced time costs and improved computational efficiency, allowing for the simulation of networks on personal computers to study the self-organized behavior of network conductivity change pathways. It is worth noting that the conduction pathways formed after

the simulation are consistent with the experimental evidence reported by Li *et al.*[28] and the computational model reported by Manning *et al.*[29]

In this article, we directly observe and analyze the network voltage distribution through the model. During the experiment, we apply a constant pulse voltage stimulation to the network voltage input terminal while maintaining the output terminal voltage at zero. As shown in Fig. 3(a), the voltage distribution of different nodes in the network under external voltage stimulation is plotted. Through this model, we can directly observe and analyze the voltage changes of network nodes at each moment. From Fig. 3(a), we can find that under constant pulse voltage stimulation, the voltage changes of different nodes in the network exhibit variability. This variability is due to the spontaneous formation and dissolution of conduction pathways caused by the memristive edge conductance characteristics of the network, resulting in different nodes in the network having different voltage changes in the time dimension. This endows the network with rich emergent dynamic behaviors.

To further demonstrate the spontaneous behavior characteristics of the network conduction pathways, we apply constant pulse voltage stimulation to 10 different pairs of terminals in the network (these 10 pairs of terminals have the same voltage input terminals and different output terminals) and record the voltage changes of the same nanowire node. Figure 3(b) shows the voltage differences and emerging dynamic behaviors of some network nodes under different terminals with constant pulse voltage stimulation.

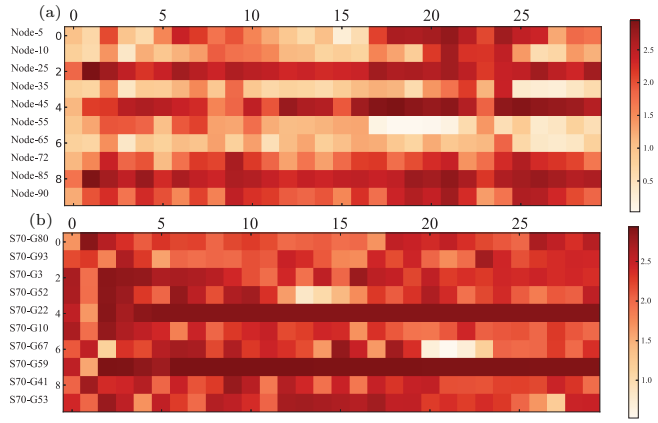


Fig. 3. (a) Simulation of the different nodes' voltage changes V in the network under a constant external voltage stimulus of 3 V. The color shades in the figure indicate the node voltage intensity. (b) Simulation of the same node voltage changes V in the network under a constant external voltage stimulus of 3 V, but with different positions of the network output terminal while keeping the voltage input terminal fixed. The color shades also represent the node voltage intensity.

The experimental results show that when constant pulse voltage stimulation is applied to different terminals, different conduction pathways will spontaneously form between network nodes, leading to different voltage changes for the same node. This phenomenon indirectly reflects the relationship between the emergent dynamic behavior inside the network and the position of external electrical stimulation. The experiment proves that, under external stimulation, the network interior will generate complex emergent dynamic behaviors. Through this model, we can analyze the network node voltage, network conductance changes, and the formation and dissolution of network conduction pathways in the time domain. Therefore, this model can serve as an effective tool for studying and learning the dynamic behavior of nanowire networks.

3. The DS nanowire network

3.1. Deep echo state network

Reservoir computing models have gained popularity in various applications, including image classification[30] and time series prediction.[31,32] The echo state network (ESN),[33] a widely used reservoir computing model, differs in function and structure from artificial neural networks and recurrent neural networks. ESNs employ a random sparse matrix to determine their internal connections. Data is mapped to a high-dimensional dynamic space by the reservoir network and then input to the readout layer for classification or prediction. In ESNs, only the readout layer weights require training, significantly reducing computational effort and enhancing the network's efficiency. The random cyclic connections in the ESN's reservoir layer are well-suited to time series prediction and have achieved notable results across various fields.[34–37] To augment the information processing capabilities of ESNs, Pedrelli *et al.*[38] introduced the deep ESN concept, which emphasizes the hierarchical nature of the reservoir. Deep ESNs offer considerable advantages over single-layer ESNs, improving the reservoir state's richness and the network's memory capacity. The deep ESN structure is depicted in Fig. 4. In this network, the first layer is internally connected and functions similarly to the reservoir layer in a single-layer ESN. As the network deepens, the previous layer's state is input to the subsequent layer.

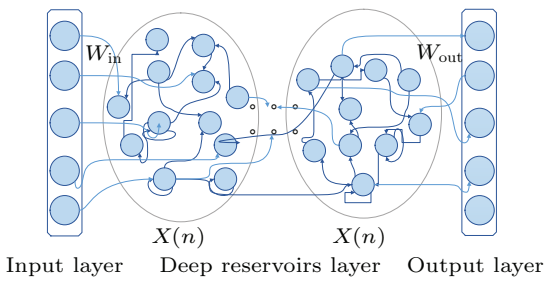


Fig. 4. Deep echo state network structure.

The first layer state transfer function is defined as follows:

$$\boldsymbol{x}_{[1]}(t) = (1-a_{[1]})\boldsymbol{x}_{[1]}(t-1)+a_{[1]}\tanh(W_{\text{in}}u(t)) + \hat{W}_{[1]}\boldsymbol{x}_{[1]}(t-1). \tag{6}$$

When the number of layers is greater than 1, the state transfer function is defined as follows:

$$\boldsymbol{x}_{[m]}(t) = (1-a_{[m]})\boldsymbol{x}_{[m]}(t-1) + a_{[m]}\tanh(W^m\boldsymbol{x}^{m-1}(t)) + \hat{W}_{[m]}\boldsymbol{x}_{[m]}(t-1), \tag{7}$$

where each network layer has the same reservoir unit N, the subscript $[m]$, denotes the mth layer of the network, time t and $t-1$ denote the current moment t and the last moment at layer m. W_{in} is the input weight matrix, $W_{[m]}$ is the interlayer connection weight matrix from layer $m-1$ to layer m, $\hat{W}_{[m]}$ is the recursive weight matrix of layer m, $a_{[m]}$ is the leakage parameter of layer l. The tanh is equivalent to the activation function in a convolutional neural network providing nonlinear states.

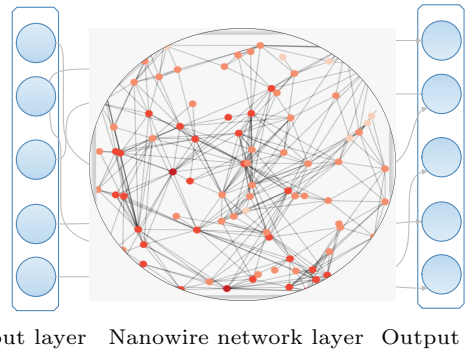


Fig. 5. A conceptual schematic of the random nanowire network-based reservoir computing for time series prediction. The steps involve (1) data preprocessing, (2) data encoding and transmission in the random nanowire network, and (3) readout layer processing and prediction of the output sequence from the reservoir.

Pedrelli *et al.*[38] introduced a deep neural network model that efficiently handles complex time series data by exploiting hierarchical temporal scale differences. Based on the deep ESN model, we propose double-layer stochastic (DS) nanowire networks to overcome the limitations of single-layer nanowire networks in generating rich nonlinear behaviors. This approach is a key step towards developing deep reservoir computing based on nanowire networks. The DS architecture enhances the predictive capability of reservoir computing based on nanowire networks for time series. As shown in Fig. 5, reservoir computing is implemented using a single-layer nanowire network with N nodes. We suggest converting the data into pulse voltages with different amplitudes and then inputting them to the voltage input terminal node. According to the principle of reservoir computing, the input voltage data is mapped to various node voltages within the nanowire network, and time series information is processed by analyzing the voltage changes of different nodes in the time domain. First, we normalize the input time series data between 0 and 1 and input it to the nanowire network. Then, we convert the data into pulse voltages with different amplitudes and introduce them to the voltage input terminal. As the input voltage fluctuates, the node voltages within the nanowire network change continuously, and the voltage change of each node depends on its historical input voltage. When an external voltage stimulus is applied to the network, the voltage distribution of different nodes in the nanowire network changes. By analyzing the voltage changes of different nodes over time, we can predict data trends. In the whole process, we read a part of the node voltage state values at each moment for subsequent data analysis and processing. Finally, we input all the obtained node voltage state values to the network readout layer for further processing and interpretation.

The role of the readout layer in the DS nanowire network is to convert the internal state of the network into usable outputs, enabling the network to make predictions, perform classifications, or other tasks. The readout layer is the only part of the network that undergoes training and is typically solved using the ridge regression method. In the DS nanowire network we designed, the global reservoir state $\boldsymbol{X}$ in the network is composed of the states from different network layers. As shown in Fig. 6(a), the readout layer is directly connected to nodes in two different nanowire network layers using a fully connected approach. This allows for direct collection of the state of each nanowire network node for training purposes, and the internal voltage state vectors of different network layers are horizontally stacked to generate a global state vector. Regarding the training algorithm, the output layer in the DS nanowire network is trained using the standard reservoir computing paradigm.[39] Specifically, we use the ridge regression algorithm to obtain the output weight matrix $\boldsymbol{W}_{\text{out}}$ as follows:

$$\boldsymbol{W}_{\text{out}} = \boldsymbol{Y}_{\text{target}}\boldsymbol{X}^{\text{T}}(\boldsymbol{X}\boldsymbol{X}^{\text{T}} + \gamma\boldsymbol{I})^{-1}. \tag{8}$$

Here, $\boldsymbol{Y}_{\text{target}}$ represents the target for prediction, $\boldsymbol{X}$ represents the high-dimensional vector generated by the nanowire network layer, $\boldsymbol{I}$ represents the identity matrix, γ is the regularization coefficient and $\boldsymbol{I}$ represents the identity matrix. The symbols "T" and "-1" denote matrix transpose and inverse operations, respectively.

To summarize, in the DS nanowire network, the global reservoir state $\boldsymbol{X}$ is composed of the states from different network layers. The output layer is trained using the ridge regression algorithm based on the standard reservoir computing paradigm to obtain the output weight matrix $\boldsymbol{W}_{\text{out}}$. During training, the output weight matrix is computed by multiplying and inverting the product of the target for prediction $\boldsymbol{Y}_{\text{target}}$, the high-dimensional vector generated by the training layer $\boldsymbol{X}$, the regularization coefficient γ, and the identity matrix $\boldsymbol{I}$. After training is completed, $\boldsymbol{W}_{\text{out}}$ can be directly used for prediction.

We propose a DS nanowire network model, which consists of two nanowire network layers with unique network connectivity and random node distribution, as depicted in

Fig. 6. This design overcomes the limitations of single-layer nanowire networks, such as limited node number, reduced network complexity and unsatisfactory predictive performance. The DS-nanowire network is primarily manifested in two aspects: first, the randomness in the connections between nanowire nodes within each layer of the multi-layer nanowire network, and second, the variability in the connections between different layers of the nanowire network. Additionally, the inter-layer connections are established through random one-to-one connections, ensuring both accurate data transmission and enhanced nonlinearity and expressive power of the network. By increasing the nanowire nodes and generating more diversified node states, the model is more suitable for time series prediction, providing an implementation strategy for reservoir computing in multi-layer nanowire networks.

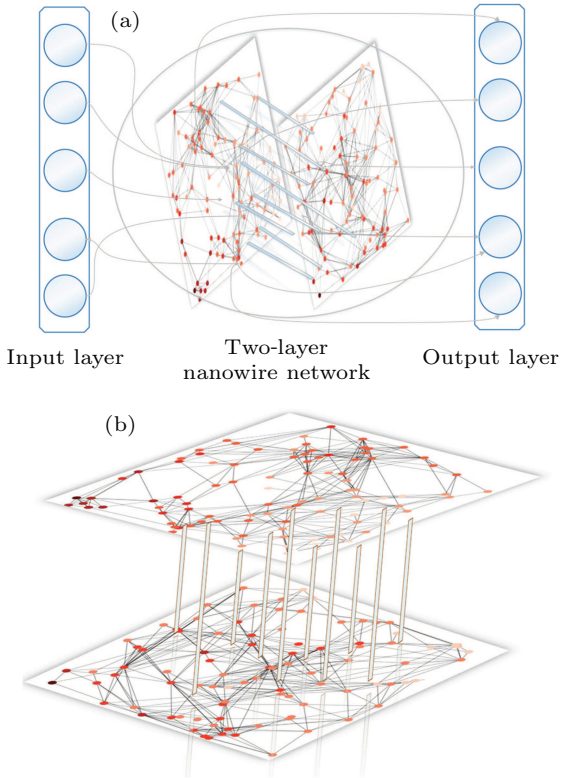


Fig. 6. (a) Conceptual schematic of the DS nanowire network-based reservoir computing for time series prediction. (b) Schematic diagram of the nanowire reservoir structure of the DS nanowire network model. The reservoir structure consists of a randomly generated dual-layer nanowire network, and the network layers are connected using a one-to-one node mapping method, i.e., each node in the upper network corresponds to a node in the lower network.

The proposed DS nanowire network model has a novel reservoir network structure with single-input-multi-output and multi-input-multi-output capabilities, which are formed by processing the node voltage state values of both network layers in the readout layer. In the DS nanowire network, the input data is mapped to intermediate node voltage state values through the first nanowire network layer at each moment, for a certain number of nodes (15 nodes are used in the experiment). Taking advantage of the ability of nanowire networks to process multiple input signals in spatiotemporal modes, the node voltage state vector serves as the input for the subsequent network layer. The lower layer network input terminal accepts the intermediate node voltage state vector as the network input, processes the intermediate node state values and outputs node voltage state values. The DS architecture is adopted for practical reasons, as it is challenging to fabricate two identical nanowire networks in the production process. The architecture is compatible with the nanowire network layer and ensures satisfactory performance. We evaluate the predictive capability of the DS nanowire network using three real-world datasets. By combining nanowire networks with reservoir computing and using node voltages as intermediate voltage states, the nanowire network structure is consistent with the reservoir layer structure of reservoir computing, paving the way for hardware implementation of deep reservoir computing.

4. Experiment

4.1. Dataset introduction

In this section, we assess the prediction performance of single-layer ESN, layered ESN, single-layer nanowire network, and our proposed DS-nanowire networks on three real-world time series datasets: the sunspot time series dataset, the Birmingham parking dataset and the bike-sharing dataset.

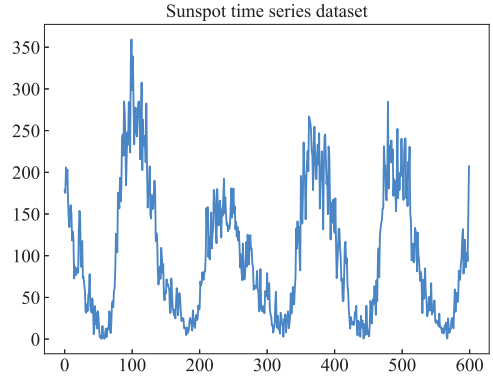


Fig. 7. Sunspot time series dataset(first 1000 points). The x-axis coordinates represent the collection sampling points and the y-axis coordinates represent the sunspot data.

Sunspot time-series predictions Sunspots reflect the dynamics of the strong magnetic field surrounding the Sun, and their numbers correlate with solar activity. The unpredictability and complexity of solar activity result in strong nonlinearities in sunspot sequences, making them suitable for testing network models' ability to solve real-world time-series prediction problems. We use a 13-month open-source monthly unsmoothed sunspot series provided by the World Data Centre SILSO,[40] collected from January 1749 to March 2021, containing 3267 sunspot sampling points. In our experiment, the first 2200 sequence sampling points serve as the network training set, 2200–2400 sequence sampling points as the validation set, and 2400–3000 sequence sampling points as the test set,

performing one-step ahead direct prediction. The sunspot time series data is shown in Fig. 7.

Birmingham parking dataset Sourced from the UCI Machine Learning Repository,[41] this dataset records occupancy rates for all council-managed car parks, collected from October 4, 2016 to December 19, 2016. The data is normalized to a 0–1 range for testing purposes. We use 10000 recent observations in our experiments, with 80% (8000) for training and 20% (2000) for testing, performing the same one-step-ahead prediction. The Birmingham parking dataset is shown in Fig. 8.

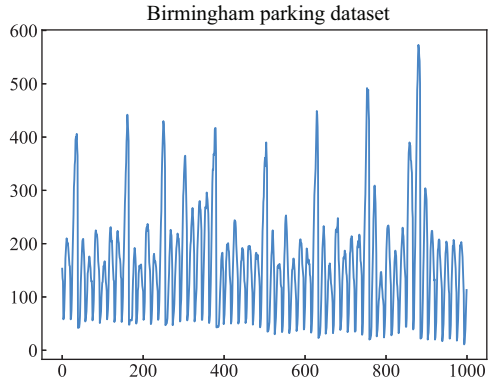


Fig. 8. Birmingham parking dataset (first 1000 points). The x-axis coordinates represent the collection sampling points and the y-axis coordinates represent the Birmingham parking data.

Bike-sharing dataset Obtained from the UCI Machine Learning Repository,[42] this dataset contains the hourly number of bicycles rented in the bike-sharing system, collected from January 1, 2011 to December 31, 2012. We use 10000 recent observations in our experiments, with 80% (8000) for training and 20% (2000) for testing, performing the same one-step-ahead prediction. The Bike-sharing dataset is shown in Fig. 9.

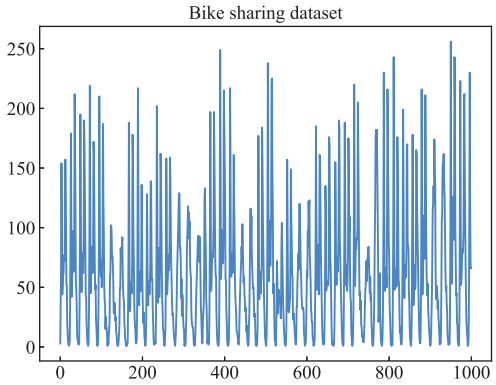


Fig. 9. Bike sharing dataset(first 1000 points). The x-axis coordinates represent the collection sampling points and the y-axis coordinates represent the Birmingham parking data.

To ensure fairness in the experiments and to highlight the advantages of the dual random nanowire networks, all single-layer ESN networks and layered ESN networks have the same number of training parameters for the output layer. The size of each reservoir layer is fixed at 300, a reasonable size that ensures both network performance and operational efficiency. All single-layer nanowire networks and layered nanowire networks have a fixed number of training parameters for the output layer of 100, with 15 nodes' voltages serving as the intermediate state node voltages for data analysis. All single-layer and layered models were compared, with the intermediate reservoir layer and output connections added to all models.

4.2. Evaluation metrics

It is important to note that there are two forecasting methods for long-term time series forecasting: iterative forecasting and direct forecasting. Iterative forecasting predicts subsequent values based on previous ones, while direct forecasting establishes a multi-step mapping relationship without iterating to obtain the result after multiple steps. The forecasting model can be expressed as follows:

$$y_{t+n} = f_n\left(y_t, \ldots, y_{t-d+1}\right) + w. \tag{9}$$

The original ESN[43] used iterative prediction, which has higher accuracy than direct prediction. However, previous works[44–46] found that iterative prediction suffers from error accumulation, leading to instability when dealing with real-world nonlinear time series. Therefore, direct prediction is used in the experiments. Normalized mean square error (NMSE), mean absolute percentage error (MAPE) and normalized root mean square error (NRMSE) are employed to evaluate the prediction performance of each model. The NRMSE formula is as follows:

$$\mathrm{NRMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(Y(i) - y(i)\right)^2}{\sum_{i=1}^{N}\left(Y(i)\right)^2}}. \tag{10}$$

The closer the evaluation result is to 0, the higher the prediction accuracy. Here, $Y_{(t)}$ represents the target signal, $y_{(t)}$ denotes the prediction result after training and i indicates the number of time points. The MAPE formula is as follows:

$$\mathrm{MAPE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{N}}\left|\frac{Y(i) - y(i)}{Y(i)}\right| \times 100\%, \tag{11}$$

where $Y(t)$ is the target signal, $y(t)$ is the prediction result after training and N denotes the number of time points. The evaluation results range from percentage results, where an MAPE of 0% indicates a perfect model and an MAPE greater than 100% indicates an inferior model. The NMSE error formula is as follows:

$$\mathrm{NMSE} = \sum_{i=1}^{N}\frac{Y(i) - y(i)^2}{Y(i)^2}. \tag{12}$$

The closer the evaluation result is to 0 indicates the higher prediction accuracy, where $Y(t)$ is the target signal, $y(t)$ is the prediction result after training and i denotes the number of times points.

4.3. Results

The performance metrics for various network models on the sunspot dataset are presented in Table 1. Additionally, Fig. 10 compares the predicted and actual results of the layered ESN and DS nanowire network for a selected time step in the sunspot dataset. The DS nanowire network's predictions are relatively smooth, with improved MAPE metrics, outperforming the layered ESN. Despite having fewer nodes than the ESN and using the node voltage of some nodes (15 nodes in this case) for model training, the DS nanowire network still achieves excellent real-time series prediction results. The DS nanowire network's prediction index on the sunspot time series dataset is higher than that of the single-layer nanowire network, which is related to the dataset's data quantity and model complexity. The DS nanowire network can overcome data quantity limitations and enhance prediction accuracy, demonstrating the rationality and superiority of a deep reservoir computing design approach based on nanowire networks for real-time series prediction.

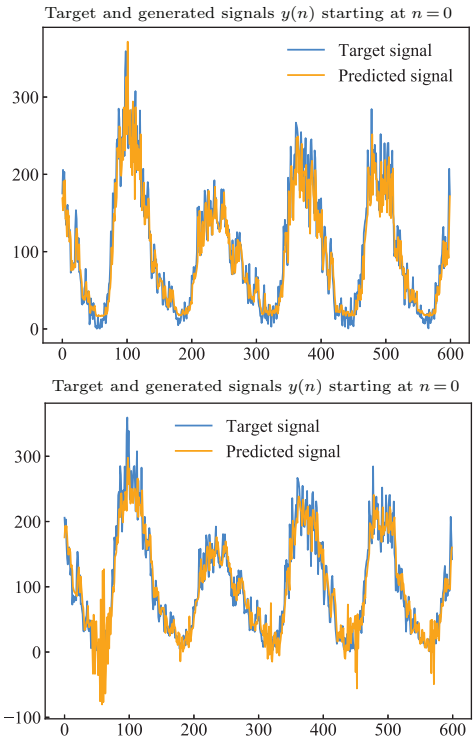


Fig. 10. (a) Plot of the effect of the predicted results of the DS nanowire network model for predicting sunspot time series compared with the actual results. (b) Plot of the effect of the predicted results of the layered ESN model for predicting sunspot time series compared with the actual results.

Furthermore, Fig. 11 compares the layered ESN and DS nanowire network's prediction results for a selected time step in the Birmingham parking dataset and the actual results. During the prediction process, the layered ESN model may exhibit significant discrepancies between predicted and actual results. In contrast, the DS nanowire network model demonstrates smooth data point transitions, with minimal occurrence of unstable data prediction points. It consistently generates stable and accurate predictions at each prediction point, outperforming the hierarchical ESN in this aspect.

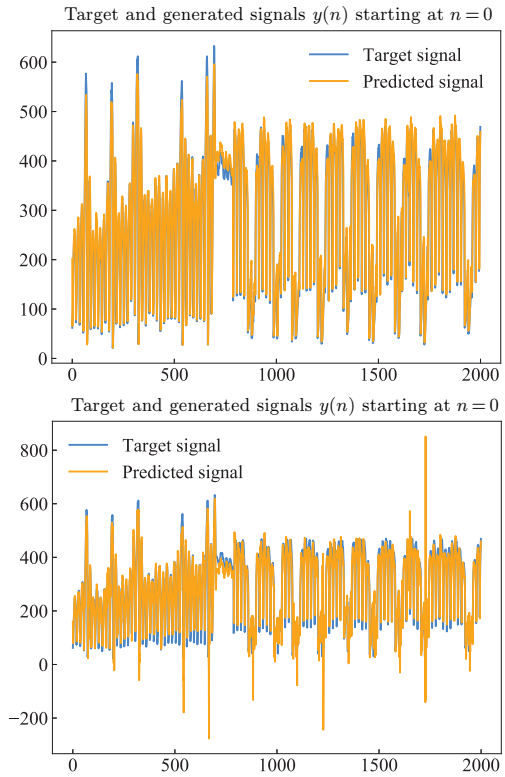


Fig. 11. (a) Plot of the predicted results of the DS nanowire network model predicting the Birmingham parking dataset compared to the actual results. (b) Plot of the predicted results of the layered ESN model predicting the Birmingham parking dataset compared to the actual results.

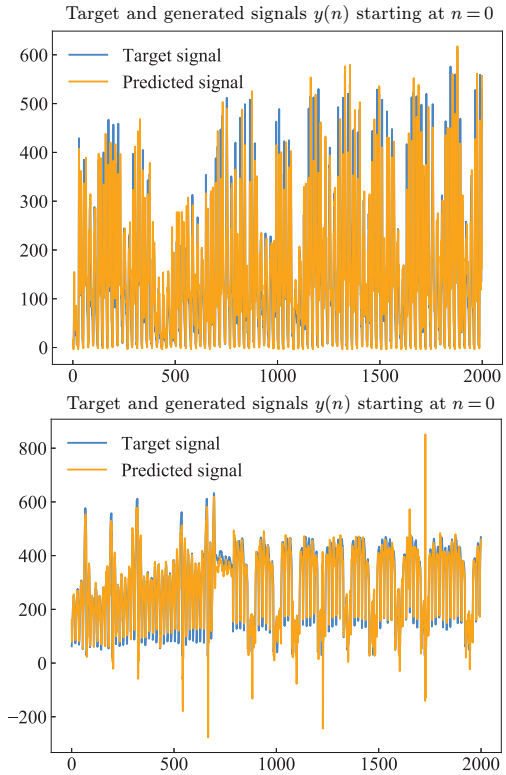


Fig. 12. (a) Plot of the predicted results of the DS nanowire network model predicting the bike-sharing dataset compared to the actual results. (b) Plot of the predicted results of the layered ESN model predicting the bike-sharing dataset compared to the actual results.

Additionally, Fig. 12 compares the layered ESN and DS nanowire network's prediction results for a selected time step in the bike-sharing dataset and the actual results. The DS nanowire network model not only achieves better accuracy but also ensures the stability of the prediction results. It possesses the flexibility to adapt to data variations, enabling the model to make accurate predictions in complex real-world environments.

According to Tables 1–3, the simulation results show that the three tables contain multiple models' prediction metrics for the sunspot dataset, the Birmingham parking dataset and bike-sharing dataset. Combined with Figs. 10–12, we will analyze the tables in detail from two aspects: prediction accuracy and data metrics analysis, and further explore the advantages of the DS nanowire network in terms of prediction stability, flexibility and robustness.

Prediction accuracy and data metrics analysis: prediction accuracy is an important metric for measuring the error between model predictions and actual observations. Based on the MAPE metrics in Tables 1–3, the DS nanowire network model demonstrates relatively high prediction accuracy across all three datasets. In Table 1, the DS nanowire network model has an MAPE of 0.685, which is lower compared to other models. In Tables 2 and 3, the DS nanowire network model has MAPE values of 0.248 and 0.873, respectively, which are also relatively low. This indicates that the DS nanowire network model can predict the target variable more accurately with smaller errors between the predicted and observed values.

Table 1. Performance of the DS nanowire network model with other prediction models on the sunspot time series.

	Single-layer ESN	Layered ESN (layers = 2)	Single-layer nanowire network	**DS nanowire network**
NARMSE	4.273×10^{-1}	3.753×10^{-1}	3.482×10^{-1}	3.359×10^{-1}
NMSE	1.826×10^{-1}	1.408×10^{-1}	1.212×10^{-1}	1.128×10^{-1}
MAPE	2.252	1.065	0.606	0.685

Table 2. Performance of the DS nanowire network model with other prediction models on the Birmingham parking dataset.

	Single-layer ESN	Layered ESN (layers = 2)	Single-layer nanowire network	**DS nanowire network**
NARMSE	5.139×10^{-1}	4.169×10^{-1}	4.076×10^{-1}	3.936×10^{-1}
NMSE	2.640×10^{-1}	1.738×10^{-1}	1.661×10^{-1}	1.549×10^{-1}
MAPE	0.337	0.311	0.249	0.248

Table 3. Performance of the DS nanowire network model with other prediction models on the bike-sharing dataset.

	Single-layer ESN	Layered ESN (layers = 2)	Single-layer nanowire network	**DS nanowire network**
NARMSE	4.590×10^{-1}	3.847×10^{-1}	4.125×10^{-1}	3.838×10^{-1}
NMSE	2.107×10^{-1}	1.480×10^{-1}	1.702×10^{-1}	1.473×10^{-1}
MAPE	1.243	1.312	0.694	0.873

Prediction stability Prediction stability measures the consistency and stability of a model across different datasets. Based on the NARMSE and NMSE metrics in Tables 1–3, the DS nanowire network model has relatively low values, indicating smaller errors between the predicted and observed values and consistent performance across different datasets. This suggests that the DS nanowire network model has an advantage in prediction stability, producing stable and accurate predictions across different datasets.

Flexibility Flexibility refers to the ability of a model to adapt to different datasets and problems. Based on the data metrics in Tables 1–3, the DS nanowire network model has lower model complexity compared to other models, making it easier to understand, implement and adjust. Its flexibility allows it to adapt to different datasets and problems, and it exhibits good generalization ability. This flexibility makes the DS nanowire network model more practical and scalable in real-world applications.

Robustness Robustness refers to a model's stability and adaptability to noise and variations in the data. Based on the prediction metrics in Tables 1–3, the DS nanowire network model demonstrates a certain level of robustness to noise and variations in the data. This means that it can generate reliable predictions in the presence of uncertainty and noise, showing good robustness. This enables the model to make accurate predictions in complex real-world environments.

The DS nanowire network model has advantages in prediction accuracy, prediction stability, flexibility and robustness. Its lower MAPE values indicate smaller errors between predicted and observed values, indicating higher prediction accuracy. Additionally, it has lower model complexity, making it easier to understand, implement and adjust, while also exhibiting good prediction stability and robustness. These advantages make the DS nanowire network model promising for practical applications and worthy of further research and exploration.

5. Conclusions

Self-organizing random nanowire networks can physically implement neuromorphic information processing and are a new approach to achieving neuromorphic computing and in-

formation processing. This study develops a model suitable for stochastic nanowire networks and observes the network's overall dynamic behavior based on model simulations of network conductance changes and node voltage distribution over time. By leveraging the nanowire network's characteristics of adaptive reallocation of node voltage and nonlinear voltage changes, information is encoded into voltages of different amplitudes and input into the nanowire network. The node voltage changes of each node over time are then collected as a reservoir state for reservoir computing. A scheme is designed to implement reservoir computing using the nanowire network for time information processing. Based on the nanowire network's ability to process multiple spatiotemporal inputs, the DS nanowire network structure is proposed, providing a practical implementation scheme for the subsequent physical realization of deep reservoir computing based on nanowires. The single-layer nanowire network and the DS nanowire network's prediction performance are tested on three real-time series datasets. The reservoir computing model based on the nanowire network achieves better prediction results than traditional deep reservoir computing.

Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant Nos. U20A20227,62076208, and 62076207), Chongqing Talent Plan "Contract System" Project (Grant No. CQYC20210302257), National Key Laboratory of Smart Vehicle Safety Technology Open Fund Project (Grant No. IVSTSKL-202309), the Chongqing Technology Innovation and Application Development Special Major Project (Grant No. CSTB2023TIAD-STX0020), College of Artificial Intelligence, Southwest University, and State Key Laboratory of Intelligent Vehicle Safety Technology.

References

[1] Upadhyay N K, Jiang H, Wang Z, Asapu S, Xia Q and Joshua Yang J 2019 *Advanced Materials Technologies* **4** 1800589
[2] Yang J J, Pickett M D, Li X, Ohlberg D A, Stewart D R and Williams R S 2008 *Nat. Nanotechnol.* **3** 429
[3] Wang Z, Joshi S, Savel'ev S E, Jiang H, Midya R, Lin P, Hu M, Ge N, Strachan J P, Li Z, *et al*. 2017 *Nat. Mater.* **16** 101
[4] Sun L, Wang Z, Jiang J, Kim Y, Joo B, Zheng S, Lee S, Yu W J, Kong B S and Yang H 2021 *Science Advances* **7** eabg1455
[5] Moon J, Ma W, Shin J H, Cai F, Du C, Lee S H and Lu W D 2019 *Nat. Electron.* **2** 480
[6] Lao J, Yan M, Tian B, Jiang C, Luo C, Xie Z, Zhu Q, Bao Z, Zhong N, Tang X, *et al*. 2022 *Advanced Science* **9** 2106092
[7] Christensen D V, Dittmann R, Linares-Barranco B, Sebastian A, Le Gallo M, Redaelli A, Slesazeck S, Mikolajick T, Spiga S, Menzel S, *et al*. 2022 *Neuromorphic Computing and Engineering* **2** 022501
[8] Xia Q and Yang J J 2019 *Nat. Mater.* **18** 309
[9] Zhu R, Hochstetter J, Loeffler A, Diaz-Alvarez A, Nakayama T, Lizier J T and Kuncic Z 2021 *Scientific Reports* **11** 1
[10] Mallinson J, Shirai S, Acharya S, Bose S, Galli E and Brown S 2019 *Science Advances* **5** eaaw8438
[11] Hochstetter J, Zhu R, Loeffler A, Diaz-Alvarez A, Nakayama T and Kuncic Z 2021 *Nat. Commun.* **12** 1
[12] Scharnhorst K S, Carbajal J P, Aguilera R C, Sandouk E J, Aono M, Stieg A Z and Gimzewski J K 2018 *Jpn. J. Appl. Phys.* **57** 03ED02
[13] Milano G, Pedretti G, Montano K, Ricci S, Hashemkhani S, Boarino L, Ielmini D and Ricciardi C 2022 *Nat. Mater.* **21** 195
[14] Fu K, Zhu R, Loeffler A, Hochstetter J, Diaz-Alvarez A, Stieg A, Gimzewski J, Nakayama T and Kuncic Z 2020 *Reservoir computing with neuromemristive nanowire networks 2020 International Joint Conference on Neural Networks (IJCNN)* (IEEE) pp. 1–8
[15] Lilak S, Woods W, Scharnhorst K, Dunham C, Teuscher C, Stieg A Z and Gimzewski J K 2021 *Frontiers in Nanotechnology* **3** 675792
[16] Schrauwen B, Verstraeten D and Van Campenhout J 2007 *An overview of reservoir computing: theory, applications and implementations Proceedings of the* 15*th european symposium on artificial neural networks* pp. 471–482 2007 pp. 471–482
[17] Milano G, Pedretti G, Fretto M, Boarino L, Benfenati F, Ielmini D, Valov I and Ricciardi C 2020 *Advanced Intelligent Systems* **2** 2000096
[18] Tanaka H, Akai-Kasaya M, TermehYousefi A, Hong L, Fu L, Tamukoh H, Tanaka D, Asai T and Ogawa T 2018 *Nat. Commun.* **9** 1
[19] Shirai S, Acharya S K, Bose S K, Mallinson J B, Galli E, Pike M D, Arnold M D and Brown S A 2020 *Network Neuroscience* **4** 432
[20] Diaz-Alvarez A, Higuchi R, Sanz-Leon P, Marcus I, Shingaya Y, Stieg A Z, Gimzewski J K, Kuncic Z and Nakayama T 2019 *Scientific Reports* **9** 1
[21] Montano K, Milano G and Ricciardi C 2022 *Neuromorphic Computing and Engineering* **2** 014007
[22] Zegarac A and Caravelli F 2019 *Europhys. Lett.* **125** 10001
[23] Milano G, Miranda E and Ricciardi C 2022 *Neural Networks* **150** 137
[24] Menzel S, Tappertzhofen S, Waser R and Valov I 2013 *Physical Chemistry Chemical Physics* **15** 6945
[25] Citri A and Malenka R C 2008 *Neuropsychopharmacology* **33** 18
[26] Rodriguez-Fernandez A, Cagli C, Suñé J and Miranda E 2018 *IEEE Electron. Device Lett.* **39** 656
[27] Miranda E, Milano G and Ricciardi C 2020 *IEEE Transactions on Nanotechnology* **19** 609
[28] Li Q, Diaz-Alvarez A, Iguchi R, Hochstetter J, Loeffler A, Zhu R, Shingaya Y, Kuncic Z, Uchida K I and Nakayama T 2020 *Advanced Functional Materials* **30** 2003679
[29] Manning H G, Niosi F, da Rocha C G, Bellew A T, O'Callaghan C, Biswas S, Flowers P F, Wiley B J, Holmes J D, Ferreira M S, *et al*. 2018 *Nat. Commun.* **9** 1
[30] Sun J, Li L and Peng H 2021 *An image classification method based on echo state network* 2021 *International Conference on Neuromorphic Computing (ICNC)* (IEEE) pp. 165–170
[31] Gao R, Du L, Duru O and Yuen K F 2021 *Applied Soft Computing* **102** 107111
[32] Li D, Han M and Wang J 2012 *IEEE Transactions on Neural Networks and Learning Systems* **23** 787
[33] Jaeger H 2002 *Advances in neural information processing systems* 15
[34] Song Z, Wu K and Shao J 2020 *Neurocomputing* **406** 343
[35] Shi G, Liu D and Wei Q 2016 *Neurocomputing* **216** 478
[36] Skowronski M D and Harris J G 2006 *Minimum mean squared error time series classification using an echo state network prediction model* 2006 *IEEE International Symposium on Circuits and Systems* (IEEE) pp. 3153–3156
[37] Lin X, Yang Z and Song Y 2009 *Expert Systems with Applications* **36** 7313
[38] Gallicchio C, Micheli A and Pedrelli L 2017 *Deep reservoir computing: A critical experimental analysis. Neurocomputing*, **268** pp. 87–99
[39] Lukoševičius M and Jaeger H 2009 *Computer science review* **3** 127
[40] Sun C, Song M, Cai D, *et al.* 2022 *A Systematic Review of Echo State Networks from Design to Application IEEE Transactions on Artificial Intelligence*, 2022. pp. 23–37
[41] Stolfi D H, Alba E and Yao X 2017 *Predicting car park occupancy rates in smart cities International Conference on Smart Cities* (Springer) pp, 107–117
[42] Fanaee-T H and Gama J 2014 *Progress in Artificial Intelligence* **2** 113
[43] Jaeger H and Haas H 2004 *Science* **304** 78
[44] Shi Z and Han M 2007 *IEEE Transactions on Neural Networks* **18** 359
[45] Sorjamaa A, Hao J, Reyhani N, Ji Y and Lendasse A 2007 *Neurocomputing* **70** 2861
[46] Xu M and Han M 2016 *IEEE Transactions on Cybernetics* **46** 2173