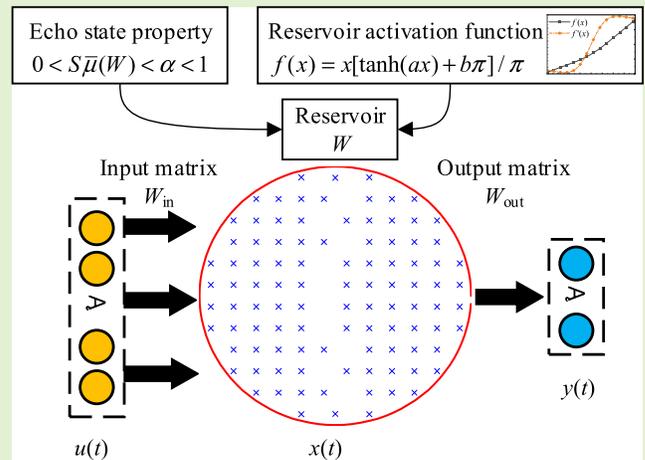


An Echo State Network With Improved Topology for Time Series Prediction

Xin Li¹, Fengrong Bi¹, Xiao Yang¹, and Xiaoyang Bi

Abstract—An echo state network with improved topology (IESN) is proposed for accurate and efficient time series prediction. In this approach, a tighter bound of the echo state property related to the Lipschitz constant of reservoir activation function and the maximum structured singular value of reservoir is firstly researched to run the model at the edge of chaos. A smooth composite reservoir activation function is then designed to enhance the ESN. The exact echo state property bound is solved by computing the Lipschitz constant of the composite function. Finally, a decoupling matrix with eigenvalues distributing uniformly in the complex plane is built as the reservoir for abundant dynamic characteristics. Six classical benchmarks are employed to test the IESN. Besides, combined with amplitude-frequency separation based on the Hilbert transform, the IESN predicts a set of engine vibration signals in knock. Compared with several popular models, the proposed IESN shows the best performance.

Index Terms—Echo state network (ESN), echo state property, reservoir, time series, vibration.



I. INTRODUCTION

TIME series prediction is crucial in many natural science and engineering technology fields. As an important application, vibration signal prediction has been widely used in the health monitoring of the machinery system [1]. The vibration signal usually has strong nonlinearity, and reconstructing the nonlinear system by measured data to predict subsequent time series is a great challenge [2]. With the development of computer theory, machine learning gradually gets gratifying nonlinear fitting capability and becomes a promising direction. Nevertheless, the accuracy and efficiency of the prediction model have always been difficult to balance.

Support vector machine (SVM) is a classical machine learning algorithm, which is also called support vector regression (SVR) when applied to time series prediction. The SVM maps data into high-dimensional space by a kernel function to implement regression analysis. The SVM's kernel

function and control parameters significantly influence the result, leading to the heuristic algorithm such as particle swarm optimization (PSO) [3] being widely employed to optimize it. Furtherly, training the SVM is a constrained optimization problem whose iterations consume much memory space and time. The large-scale sample set is also a great challenge for the SVM, so the artificial neural network (ANN) is introduced as another powerful tool. Multilayer perceptron (MLP) is a typical feed-forward ANN used for time series prediction. However, its simple frame is difficult to capture the subtle feature accurately, and the multistage hybrid model [4] is a practical improvement. The complicated training promotes the extreme learning machine (ELM) as an alternative [5], whose main characteristic is part of the weights in the model are generated randomly and fixed in training. The ELM plays a role in forecasting sequence [6] because of the high efficiency, whereas feed-forward ANNs have unavoidable defects for the lack of memory. The recurrent neural network (RNN) recurses sequence data by chain-connected nodes to obtain the short-term memory capacity and is especially suitable for time series analysis. Elman neural network (ENN) is a classical type of RNN [7], which uses delay operators to record the output of the hidden layer in the last time step. In applications, the two-stage structure [8] can often lead to an accurate ENN model. However, the ENN is disturbed by gradient vanishing and explosion in long-term time series prediction. As developments of the RNN, long short-term memory (LSTM) and gated recurrent unit (GRU) adopt a gating mechanism to alleviate the long-term dependence problem [9].

Manuscript received January 6, 2022; accepted January 29, 2022. Date of publication February 2, 2022; date of current version March 14, 2022. This work was supported by the Project WCDL-GH-2021-0017 funded by the State Key Laboratory of Engine Reliability, China. The associate editor coordinating the review of this article and approving it for publication was Dr. Peng Wang. (Corresponding author: Fengrong Bi.)

Xin Li, Fengrong Bi, and Xiao Yang are with the State Key Laboratory of Engines, Tianjin University, Tianjin 300350, China (e-mail: nvh_lixin@tju.edu.cn; fr_bi@tju.edu.cn; yangxiao@tju.edu.cn).

Xiaoyang Bi is with the School of Mechanical Engineering, Hebei University of Technology, Tianjin 300401, China (e-mail: xy_bi@hebut.edu.cn).

Digital Object Identifier 10.1109/JSEN.2022.3148742

The two models have been widely employed in engineering signal prediction, such as monitoring degradation trends of the gears [10], [11] and bearings [12]. Signal processing algorithms are often combined with the RNN to improve complex time series prediction accuracy. Recent studies indicate the time-varying filter-based empirical mode decomposition (TVF-EMD)-ENN [13], ensemble EMD (EEMD)-LSTM [14], and variational mode decomposition (VMD)-GRU [15] are feasible methods. However, the above ANNs are trained by the gradient descent algorithm, which requires extensive training data for convergence and is prone to local optimum. Furthermore, the backpropagation (BP) algorithm, especially the BP through time (BPTT) in RNNs, leads to an inefficient training process.

As a unique type of the RNN, the echo state network (ESN) [16] takes a randomly generated reservoir, instead of neurons, as the basic processing unit. Besides, only output weights are trained by simple linear regression to avoid local optimum and improve efficiency. These advantages of structure and training process promote its applications for time series prediction in various fields [17], [18], whereas the research of the ESN is not yet mature. Unlike traditional ANNs, the ESN is built by a reservoir with special dynamic characteristics, which is difficult to optimize by various modules. The randomly generated topology causes uncontrollable dynamics, and its design is still in dispute. Due to the complicated dynamics of the ESN, combining the ESN with other methods, such as employing a heuristic algorithm to optimize hyperparameters [19] or taking another ANN as the pretreatment model [20], is a popular solution. However, additional algorithms can not improve the ESN topology fundamentally; furthermore, complex structures negatively affect the prediction efficiency. Based on the comprehensive analysis of Jaeger *et al.* [21], the reservoir shows a crucial role in memory capacity and generalization. Therefore, instead of the randomly generated reservoir, Wang *et al.* [22] employed intrinsic plasticity (IP) algorithm to adjust reservoir weights for information maximization, and Qiao *et al.* [23] proposed an ESN with the growing reservoir (GESN) by singular value spectrum. The IP-ESN and GESN use the neuro-evolution algorithm to search for the best reservoir with abundant dynamic characteristics, whereas they lose ESN's original advantage of simple training. Designing an improved reservoir based on the dynamic characteristic analysis is an alternative. Gallicchio *et al.* [24] analyzed the reservoir by the entropy of recurrent unit activations to build a stacked-layers ESN. For compact structure, Zhang *et al.* [25] developed the decoupling method to design a simple reservoir structure. Besides reservoir, Buehner *et al.* [26] derived a tighter bound for the echo state property, and Zhao *et al.* [27] used ridge regression to train the ESN. However, the simple reservoir structure is not universal because it has strict requirements for operating conditions, which is difficult to achieve in practice. The tighter bound of the echo state property is not applicable for the most widely used leaky integrator ESN [21] and does not take the reservoir activation function into account. Based on the above analysis, current researches lack universality and systematicness.

Developing the fundamental topology to improve the accuracy of the ESN under high training efficiency is essential for time series prediction. An ESN with improved topology (IESN) is proposed based on optimizations of echo state property bound, reservoir activation function, and reservoir structure. The main contributions are as follows:

(1) A universal tighter echo state property bound is proposed to run the leaky integrator ESN at the edge of chaos.

(2) A composite reservoir activation function is designed to enhance the ESN. An exact echo state property bound is then computed based on the function.

(3) A reservoir with abundant dynamic characteristics is designed by uniformly distributing eigenvalues in the complex plane, and it has an efficient building process due to the non-iterative method.

(4) The amplitude-frequency separation based on the Hilbert transform is introduced to improve the accuracy of the IESN in high-frequency vibration signal prediction.

This paper is organized as follows: Section I introduces the research background and significance. Section II describes the fundamental theory of the ESN. The echo state property bound, reservoir activation function, and reservoir structure are optimized in Section III. In Section IV, the IESN is tested by six benchmarks and a set of engine vibration signals. The IESN is further analyzed and discussed in Section V. Conclusion and outlook are given in Section VI.

II. FUNDAMENTAL THEORY

The ESN is also described as reservoir computing because it takes a randomly generated reservoir as the basic processing unit. The reservoir is a sparse matrix activated into rich internal states to describe input signals by linear combination. In training, only output weights are adjusted by linear regression, avoiding local optimum and improving efficiency.

Suppose $u = \{u_1, u_2, \dots, u_n\}$ is the n -dimensional input signal, $x = \{x_1, x_2, \dots, x_N\}$ the internal state, and $y = \{y_1, y_2, \dots, y_m\}$ the m -dimensional output signal. The internal state at the $(t + 1)$ -th time step is:

$$x(t + 1) = f(W_{in}u(t + 1) + Wx(t) + W_{fb}y(t)) \quad (1)$$

where W_{in} is the input weight matrix, W the reservoir, W_{fb} the feedback weight matrix, and $f(\cdot)$ the reservoir activation function.

The leaky integrator ESN is the most frequently used variant:

$$x(t + 1) = (1 - \alpha\gamma)x(t) + \gamma f(W_{in}u(t + 1) + Wx(t) + W_{fb}y(t)) \quad (2)$$

where α is leaky rate and γ the gain. In general, $\gamma = 1$ and $W_{fb} = 0$ [21]:

$$x(t + 1) = (1 - \alpha)x(t) + f(W_{in}u(t + 1) + Wx(t)) \quad (3)$$

The output is:

$$y(t) = g(W_{out}[x(t); u(t)]) \quad (4)$$

where W_{out} is the output weight matrix, and $g(\cdot)$ the output activation function.

The W_{out} is trained by linear regression, and its objective function $L(\cdot)$ is:

$$L(\widehat{W}_{\text{out}}) = \left\| g^{-1}(y) - W_{\text{out}}[x; u] \right\|_2^2 \quad (5)$$

where $\|\cdot\|_2$ is the L_2 -norm and $g^{-1}(\cdot)$ the inverse function of $g(\cdot)$.

The estimated output weight matrix \widehat{W}_{out} could be solved by:

$$\widehat{W}_{\text{out}} = g^{-1}(y)[x; u]^{\Psi} = g^{-1}(y)([x; u]^{\text{T}}[x; u])^{-1}[x; u]^{\text{T}} \quad (6)$$

where the superscripts Ψ and T represent the pseudo inverse and the transpose of the matrix, respectively.

The echo state property is the base of precise results in the ESN, and it is usually ensured by the condition [21]:

$$\tilde{\rho} = \rho[W + (1 - \alpha)I] < 1 \quad (7)$$

where $\tilde{\rho}$ is named as the effective spectral radius. ρ is the spectral radius, i.e., the maximum value of eigenvalues' magnitudes and I the identity matrix.

III. ECHO STATE NETWORK WITH IMPROVED TOPOLOGY

The echo state property, reservoir activation function, and reservoir structure are optimized in this section. On this basis, an ESN with improved topology (IESN) is proposed.

A. Tighter Bound of Echo State Property

The echo state property requires a reservoir with asymptotic stability. Suppose x_t and \tilde{x}_t are different internal states and $y_t = x_t - \tilde{x}_t$. The ESN has the echo state property if $\lim_{t \rightarrow \infty} \|y_t\| = 0$ for all right infinite input sequences $u^{+\infty}$. Running the ESN at the edge of chaos could obtain high prediction accuracy. Whereas (7) is a conservative sufficient condition, and the echo state property may also be achieved by $\tilde{\rho} > 1$ [28]. Besides, it is not applicable for various reservoir activation functions. Therefore, a tighter bound of echo state property with universality for the leaky integrator ESN is proposed.

In particular, the theorem that all finite-dimensional norms are equivalent in the Banach space is the basis and frequently used in the following derivation. Suppose $\|\cdot\|_D$ is a operate norm and the operate D is a matrix with a specific structure in the field \bar{F} . The echo state property condition can be transformed as $\lim_{t \rightarrow \infty} \|y_t\|_D = 0$.

$$\begin{aligned} \|y_{t+1}\|_D &= \|x_{t+1} - \tilde{x}_{t+1}\|_D \\ &= \|(1 - \alpha)x_t + f(W_{\text{in}}u_{t+1} + Wx_t) - \dots \\ &\quad [(1 - \alpha)\tilde{x}_t + f(W_{\text{in}}u_{t+1} + W\tilde{x}_t)]\|_D \\ &= \|(1 - \alpha)(x_t - \tilde{x}_t) + [f(W_{\text{in}}u_{t+1} + Wx_t) \\ &\quad - f(W_{\text{in}}u_{t+1} + W\tilde{x}_t)]\|_D \\ &\leq (1 - \alpha) \|x_t - \tilde{x}_t\|_D + \|f(W_{\text{in}}u_{t+1} + Wx_t) \\ &\quad - f(W_{\text{in}}u_{t+1} + W\tilde{x}_t)\|_D \end{aligned} \quad (8)$$

Based on the Lipschitz continuity, there is a Lipschitz constant S making $|f(x_i) - f(x_j)| \leq S|x_i - x_j|$, $\forall x \in \mathbb{R}$.

Therefore, (8) is derived to:

$$\begin{aligned} \|y_{t+1}\|_D &\leq (1 - \alpha) \|y_t\|_D \\ &\quad + S \|(W_{\text{in}}u_{t+1} + Wx_t) - (W_{\text{in}}u_{t+1} + W\tilde{x}_t)\|_D \\ &= (1 - \alpha) \|y_t\|_D + S \|Wx_t - W\tilde{x}_t\|_D \\ &= (1 - \alpha) \|y_t\|_D + S \|Wy_t\|_D \\ &\leq (1 - \alpha) \|y_t\|_D + S \|W\|_D \|y_t\|_D \end{aligned} \quad (9)$$

Suppose z ($z \neq 0$) is an N -dimensional vector that $z = D^{-1}v$. The operate norm $\|\cdot\|_D$ is derived to:

$$\begin{aligned} \|W\|_D &= \sup_{z \neq 0} \frac{\|Wz\|_D}{\|z\|_D} = \sup_{z \neq 0} \frac{\|DWz\|_2}{\|Dz\|_2} = \sup_{v \neq 0} \frac{\|DWD^{-1}v\|_2}{\|v\|_2} \\ &= \bar{\sigma}(DWD^{-1}) \end{aligned} \quad (10)$$

where $\bar{\sigma}$ represents the maximal singular value and sup the supremum. The D is a nonsingular matrix, so $v \neq 0$.

Equation (9) can be derived to:

$$\begin{aligned} \|y_{t+1}\|_D &\leq (1 - \alpha) \|y_t\|_D + S\bar{\sigma}(DWD^{-1}) \|y_t\|_D \\ &= [1 - \alpha + S\bar{\sigma}(DWD^{-1})] \|y_t\|_D \end{aligned} \quad (11)$$

The echo state property bound is equivalent to:

$$\begin{aligned} \inf_{D \in \bar{F}} [1 - \alpha + S\bar{\sigma}(DWD^{-1})] < 1 &\Leftrightarrow 1 - \alpha + S \inf_{D \in \bar{F}} [\bar{\sigma}(DWD^{-1})] \\ &< 1 \\ &\Leftrightarrow S \inf_{D \in \bar{F}} [\bar{\sigma}(DWD^{-1})] < \alpha \end{aligned} \quad (12)$$

where inf represents the infimum.

The $\inf_{D \in \bar{F}} [\bar{\sigma}(DWD^{-1})]$ has an exact value if the W is a normal matrix [26]:

$$\inf_{D \in \bar{F}} [\bar{\sigma}(DWD^{-1})] = \bar{\mu}(W) \quad (13)$$

where the $\bar{\mu}$ is the maximum structured singular value.

In general, $0 < \alpha < 1$ and $S, \bar{\mu}(W) > 0$. The tighter bound of echo state property with universality for the leaky integrator ESN is:

$$0 < S\bar{\mu}(W) < \alpha < 1 \quad (14)$$

B. Composite Reservoir Activation Function

The most frequent activation functions in the ESN are the sigmoid and the hyperbolic tangent (tanh) functions, but they limit the model performance to a certain extent. Although the ReLU [29], Swish [30], and wavelet [31] functions are alternatives, they do not perform well in this model (the analysis is shown in Section V-C). The reason is that they are designed for ANNs trained by the BP algorithm, and convenience for the derivative is one of the crucial considerations. The training of the ESN is just a linear regression and does not involve the reservoir. Therefore, reservoir activation function can be enhanced without regard for complexity. A smooth composite function is proposed:

$$f(x) = x[\tanh(ax) + b\pi]/\pi \quad (15)$$

where $a, b \in \mathbb{R}^+$ are control parameters.

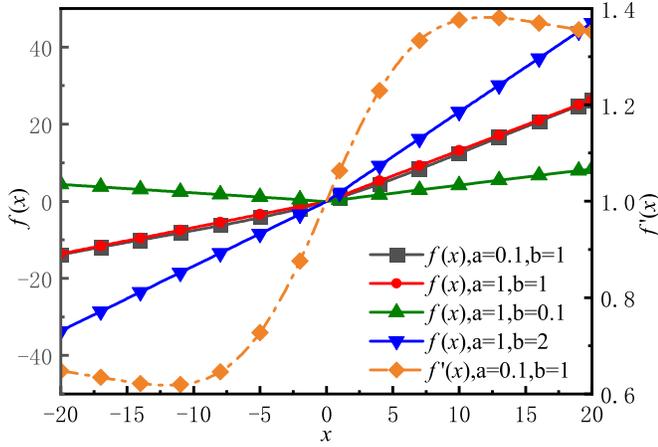


Fig. 1. Graph of the composite function.

The function graph is shown in Fig. 1, and the control parameter b has a more considerable influence on the curve shape. The tighter echo state property bound relates to the reservoir W , the leaky rate α , and the activation function $f(\bullet)$. The Lipschitz constant S is the maximum value of the reservoir activation function's derivative $f'(\cdot)$:

$$f'(x) = [\tanh(ax) + b\pi]/\pi - [ax(\tanh^2(ax) - 1)]/\pi \quad (16)$$

The limits of the $f'(x)$ can easily be solved: $\lim_{x \rightarrow +\infty} f'(x) = b + 1/\pi$ and $\lim_{x \rightarrow -\infty} f'(x) = b - 1/\pi$. However, the $f'(x)$ is nonmonotonic, as shown in Fig. 1. Its extremum values, i.e., zero-crossing points of the second derivative $f''(x)$, should be considered.

$$f''(x) = [2a(ax \tanh(ax) - 1)(\tanh^2(ax) - 1)]/\pi \quad (17)$$

The zero-crossing points are x values at $ax \tanh(ax) = 1$, which is solved as $ax \approx 1.20$. The extremum values of the $f'(x)$ could easily be computed: $b \pm 1.20/\pi$. Based on this, the Lipschitz constant S is:

$$S = b + 1.20/\pi \quad (18)$$

C. Reservoir Structure Design

A simple reservoir structure is designed in this section to obtain an exact maximum structured singular value and abundant dynamic characteristics.

Suppose $W_{\rho=1}$ is a matrix with the spectral radius $\rho = 1$. Its eigenvalues distribute in a unit circle around the origin at the complex plane and are symmetric against the real axis:

$$\lambda(W_{\rho=1}) = \{\lambda_1 + A_1i, \lambda_1 - A_1i, \lambda_2, \lambda_3 + A_3i, \lambda_3 - A_3i, \dots\} \quad (19)$$

There is a normal matrix:

$$\tilde{W}_{\rho=1} = \begin{bmatrix} \lambda_1 & -A_1 & 0 & 0 & 0 & \dots \\ A_1 & \lambda_1 & 0 & 0 & 0 & \dots \\ 0 & 0 & \lambda_2 & 0 & 0 & \dots \\ 0 & 0 & 0 & \lambda_3 & -A_3 & \dots \\ 0 & 0 & 0 & A_3 & \lambda_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (20)$$

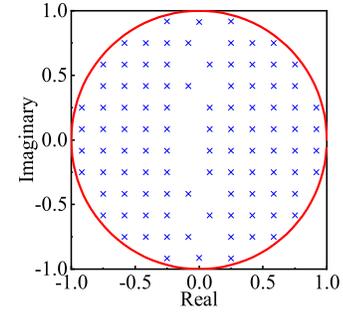


Fig. 2. Eigenvalues of the reservoir ($N = 100$) at the complex plane.

The normal matrix $\tilde{W}_{\rho=1}$ has the same eigenvalues as the $W_{\rho=1}$. The most points in $\tilde{W}_{\rho=1}$ appear in pairs, so it is described as a decoupling matrix. Reservoirs with the same eigenvalues have a similar performance [25]; therefore, the reservoir structure design is transformed as eigenvalues distributing.

Suppose the size of the $W_{\rho=1}$ is $N \times N$, so N sets of points distributing uniformly in the unit circle at the complex plane are generated for abundant dynamic characteristics. A square area around the origin with the boundary length of 2 is introduced as the auxiliary reference. The number of points in the square area is $4N/\pi$, so the number of points in every edge is $2(\sqrt{N/\pi})$. It is rounded up as $2\lceil\sqrt{N/\pi}\rceil$ to guarantee the symmetry and enough points for subsequent selection, then a square area with $4\lceil\sqrt{N/\pi}\rceil^2$ points is obtained, where the $\lceil\bullet\rceil$ represents rounding up. Pairs of points with magnitudes exceeding one are rejected firstly. The points near the imaginary axis are rejected then until the number of the rest points is N .

In particular, the number of points in the second rejection is an odd number when the N is an odd number, which will destroy the symmetry of the rest points. Although a normal matrix can still be obtained by re-decoupling, the N is usually set as an even number for simplicity. Eigenvalues ($N = 100$) acquired by this method are shown in Fig. 2.

Suppose the matrix with the designed eigenvalues is $\hat{W}_{\rho=1}$, the reservoir can be finally built based on (14):

$$Res = mu \cdot \hat{W}_{\rho=1} / \bar{\mu}(\hat{W}_{\rho=1}) \quad (21)$$

where mu is the parameter that controls the maximum structured singular value of the reservoir.

The **Algorithm** shows the design of the reservoir structure.

D. Structure of the IESN

On these bases, the structure of the IESN is shown in Fig.3. The input matrix is generated randomly, and the reservoir is designed by the **Algorithm** and activated by the (15). Then, the model could be applied through (3), (4), and (6). The signal is sent into the reservoir after the input matrix, and it is iterated based on the time step. The internal state could reach stability by certain iterations because of the echo state property. The trained output matrix could finally transform the internal state as the expected result.

Algorithm : Design the Reservoir Structure

Input: Reservoir size N and the maximum structured singular value μ .

Output: Reservoir Res .

- 1: Generate $4 \lceil \sqrt{N/\pi} \rceil^2$ points uniformly in a square area, which is around the origin with the boundary length of 2 at the complex plane.
- 2: **If** the magnitude of a point > 1 **do**
- 3: Reject the point;
- 4: **End**
- 5: **If** the number of the rest points $> N$ **do**
- 6: Reject the point near the imaginary axis in pairs until the rest number is N ;
- 7: **End**
- 8: Generate a matrix $\hat{W}_{\rho=1}$ with the rest points by (20).
- 9: Generate the reservoir Res by (21).
10. **Return** Res .

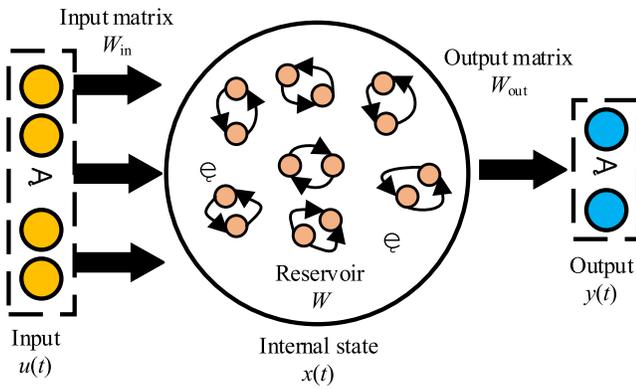


Fig. 3. Structure of the IESN.

IV. EXPERIMENT AND RESULT

The proposed IESN is tested by six classical benchmarks and a set of engine vibration signals in this section.

A. Testing Cases

1) Nonlinear Auto-Regressive Moving Average (NARMA):

The NARMA is usually used to describe a nonlinear time invariable system. A 10-order NARMA is built:

$$Y(t+1) = 0.3Y(t) + 0.05Y(t) \sum_{i=0}^9 Y(t-i) + 1.5X(t-9)X(t) + 0.1 \quad (22)$$

where $X(t)$ is a random number between 0 and 0.5.

In this case, the $Y(t)$ is used to predict the $Y(t+10)$.

2) *Mackey and Glass Time Series*: It is a nonlinear delay differential equation. The used form is:

$$dX/dt = (0.2X(t-\tau)) / (1 + X^{10}(t-\tau)) - 0.1X(t) \quad (23)$$

where $\tau = 17$.

In this case, the $X(t)$ is used to predict the $X(t+3)$.

3) *Henon Map*: The Henon map is a classical chaotic attractor:

$$\begin{cases} X(t+1) = 1 - 1.4X^2(t) + Y(t) \\ Y(t+1) = 0.3X(t) \end{cases} \quad (24)$$

In this case, the initial system values are set as $X(0) = 0$ and $Y(0) = 0$. The $X(t)$ and the $Y(t)$ are used to predict the $X(t+1)$.

4) *Lorenz Attractor*: The Lorenz attractor is the fractal structure of the Lorenz oscillator's long-term behavior. The differential equations are:

$$\begin{cases} dX/dt = 10(Y - X) \\ dY/dt = 28X - Y - XZ \\ dZ/dt = XY - (8/3)Z \end{cases} \quad (25)$$

In this case, the initial system values are set as $X(0) = -1$, $Y(0) = 0$, and $Z(0) = 1$. The $Y(t-20)$, $Y(t-10)$, and $Y(t)$ are used to predict the $X(t+10)$.

5) *Rosler Attractor*: The Rosler attractor is a simple nonlinear ordinary differential equation:

$$\begin{cases} dX/dt = -(Y + Z) \\ dY/dt = X + 0.1Y \\ dZ/dt = 0.1 + Z(X - 14) \end{cases} \quad (26)$$

In this case, the initial system values are set as $X(0) = -1$, $Y(0) = 0$, and $Z(0) = 1$. The $Y(t-10)$, $Y(t-5)$, and $Y(t)$ are used to predict the $X(t+5)$.

6) *Chen Attractor*: The Chen attractor has more complex topology and dynamic characteristics compared with the Lorenz attractor. The typical form is:

$$\begin{cases} dX/dt = 40(Y - X) \\ dY/dt = 28Y - 12X - XZ \\ dZ/dt = XY - 3Z \end{cases} \quad (27)$$

This case is tested in the same manner as the Rosler attractor case.

Actual time series are often difficult to collect, so the performance of the model with limited training data is analyzed. In the above six benchmarks, 2000 samples are generated, in which 1000 samples are taken as the training dataset, and the rest are the testing dataset.

7) *Engine Vibration*: The engine is a highly integrated mechanical system with many vibration sources and complex transfer functions. The block vibration signals contain rich information about the engine's performance and can reflect its health condition [32]. Therefore, vibration prediction for engine faults monitoring is a promising engineering application.

A four-cylinder gasoline engine tested in the knock condition, which is a common abnormal combustion phenomenon. The engine is rigidly fixed on a horizontal platform and controlled by an electric dynamometer through a carbon-fiber drive shaft, as shown in Fig. 4(a). The sensor set on the engine block is the piezoelectric accelerometer ICP 621B40 produced by PCB Piezotronics, Inc., as shown in Fig. 4(b). The knock condition is realized by advancing the ignition timing. Considering the knock feature frequency is up to 25 kHz, the sampling rate is 51200 points per second [32].

Two vibration signals near the third cylinder at 1600r/min are selected from different cycles to train and test the IESN separately for practicability. The signal during an engine cycle

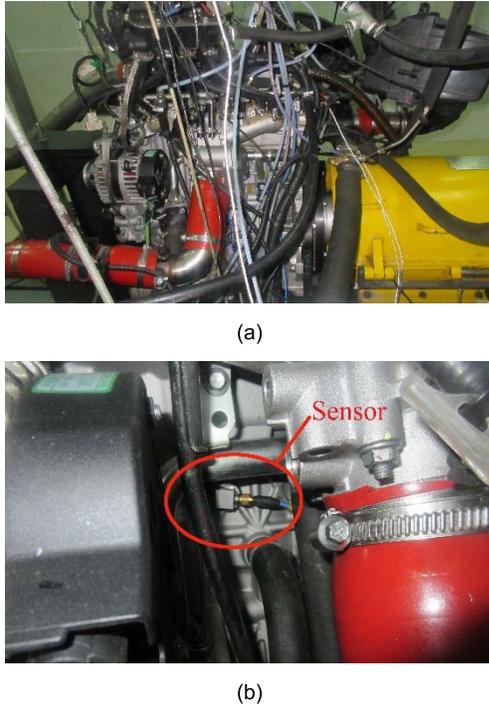


Fig. 4. Test bench. (a) Testing engine. (b) Acceleration sensor.

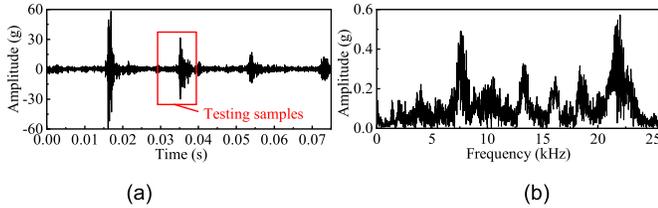


Fig. 5. Engine vibration acceleration signal. (a) Time-domain signal. (b) Frequency-domain signal.

contains 3840 samples. The knock can be detected when the signal covers the combustion period. Therefore, the 1st-2000th samples in the first signal are used to train the model, and the 1501st-2000th samples in the second one are used for testing. The time and frequency domains of the testing signal are shown in Fig. 5, where $g = 9.8\text{m/s}^2$.

In this case, time-domain signals $Y(t-40)$, $Y(t-30)$, $Y(t-20)$, $Y(t-10)$, and $Y(t)$ are used to predict the $Y(t+10)$.

B. Testing Results

Seven cases mentioned above are employed to test the IESN. Several widely used models, including the SVM, the MLP, the ELM, the ENN, the LSTM, the GRU, the original ESN, the IP-ESN, and the GESN, are tested for comparisons. Considering decomposition algorithms are effective in the complex signal analysis, three optimized RNN models, the TVF-EMD-ENN, the EEMD-LSTM, and the VMD-GRU, are also tested. The accuracy is measured by normalized root mean square error (NRMSE):

$$\text{NRMSE} = \sqrt{\frac{\sum_{t=1}^n (Y_t^{\text{predicted}} - Y_t^{\text{expected}})^2}{n \cdot \text{var}(Y^{\text{expected}})}} \quad (28)$$

where n and $\text{var}(Y^{\text{expected}})$ represent the length and the variance of the expected output, respectively.

The reservoir activation function of the IESN is $f(x) = x[\tanh(0.1x) + \pi]/\pi$, whose Lipschitz constant is $S \approx 1.38$. Therefore, the structured singular value of the reservoir is set as $\mu = 0.6$ and the leaky rate as $a = 0.928 (1.38 \times 0.6 + 0.1)$ to run the model at the edge of chaos. The reservoir size is set as $N = 100$. The SVR takes a radial basis function as the kernel. The MLP, the ENN, the LSTM, and the GRU all have three-layers structures. The VMD decomposes signals into three components. The TVF-EMD and EEMD decompose signals by stop conditions, and the results are also restructured into three components separately as inputs. The ELM has a traditional single hidden layer structure [7]. The activation function in the original ESN is the tanh, and the reservoir is generated randomly. The hyperparameters are set as $\tilde{\rho} = 0.9$ and $N = 100$ based on (7). In the IP-ESN, the reservoir is further optimized by the IP algorithm and the output matrix is solved by ridge regression. In the GESN, a sub-reservoir of 10×10 is taken to generate the reservoir. The first 100 samples are used to wash out initial conditions in the four kinds of ESN models. Considering that ANNs are sensitive to initial weights, the average values of ten runs are taken as the final results of these models (the significance analysis is shown in Section V-A).

1) *Testing Results of Benchmarks*: The testing results of six benchmarks mentioned above are listed in Table I.

As shown in Table I, the IESN has an obvious advantage, which shows the optimizations are effective. Three optimized RNN models obtain better results than the original ones in most cases, indicating that decomposition algorithms can improve the complex time series prediction to a certain extent. However, the Henon map case uses two different variables as the input, and decomposition algorithms destroy their internal relation. Therefore, the decomposition algorithm is not suitable for all applications. The IP-ESN obtains the second accuracies only to the IESN except in the Henon map case, which shows that training the ESN by the heuristic algorithm is a potential method. Nevertheless, the IESN trained by linear regression has higher efficiency and is more practical.

2) *Testing Results of Vibration Signal*: The knock is usually detected by the energy level of the high-frequency component (6-25 kHz) in the vibration signal, so the frequency-domain prediction is equally vital in this case. However, the time series prediction result and its Fourier spectra are not satisfying, as shown in Fig. 6. The NRMSEs are 1.05 in time-domain and 1.04 in frequency-domain, respectively. The high-frequency components in 20-25 kHz and 8-10 kHz are almost entirely lost. Especially the energy differences of components in 20-25 kHz are significant.

The common shortcoming of ANNs in high-frequency signal prediction is the main reason, so an amplitude-frequency separation based on the Hilbert transform is employed. Suppose the input signal is $u(t)$, and its analytic signal can be solved by the Hilbert transform:

$$u_A(t) = \frac{1}{\pi} p.v. \int_R \frac{u(v)}{t-v} dv \quad (29)$$

where $p.v.$ represents the Cauchy principal value.

TABLE I
NRMSEs OF BENCHMARK TESTING RESULTS

Models	NARMA	Mackey-Glass	Henon map	Lorenz attractor	Chen attractor	Rosslor attractor
SVM	1.01	7.36×10^{-1}	1.04×10^{-1}	3.79×10^{-1}	5.22×10^{-2}	1.75×10^{-2}
MLP	8.72×10^{-1}	8.01×10^{-1}	1.20×10^{-2}	5.12×10^{-1}	3.01×10^{-2}	8.01×10^{-1}
ELM	1.05	7.30×10^{-1}	7.82×10^{-5}	2.83×10^{-1}	1.97×10^{-1}	1.12×10^{-1}
ENN	1.00	7.32×10^{-1}	2.61×10^{-2}	2.63×10^{-1}	1.25×10^{-1}	3.12×10^{-2}
TVF-EMD-ENN	9.35×10^{-1}	2.89×10^{-1}	1.02	2.30×10^{-1}	1.18×10^{-1}	2.58×10^{-1}
LSTM	1.05	1.16×10^{-1}	8.78×10^{-2}	3.67×10^{-1}	2.14×10^{-1}	1.82×10^{-1}
EEMD-LSTM	9.72×10^{-1}	1.03×10^{-1}	1.04	3.04×10^{-1}	3.52×10^{-1}	1.56×10^{-1}
GRU	1.11	2.11×10^{-1}	4.06×10^{-2}	1.63×10^{-1}	8.33×10^{-2}	7.64×10^{-2}
VMD-GRU	9.17×10^{-1}	2.07×10^{-1}	1.05	1.00×10^{-1}	2.34×10^{-1}	3.29×10^{-2}
Original ESN	9.22×10^{-1}	1.17×10^{-1}	9.41×10^{-2}	1.65×10^{-1}	1.30×10^{-1}	5.43×10^{-2}
IP-ESN	4.42×10^{-1}	1.54×10^{-1}	9.94×10^{-2}	2.63×10^{-2}	1.85×10^{-2}	5.15×10^{-3}
GESN	6.20×10^{-1}	7.60×10^{-2}	9.09×10^{-4}	1.64×10^{-1}	1.22×10^{-1}	5.33×10^{-2}
IESN	1.90×10^{-1}	3.19×10^{-2}	2.02×10^{-5}	1.28×10^{-2}	1.30×10^{-2}	1.07×10^{-3}

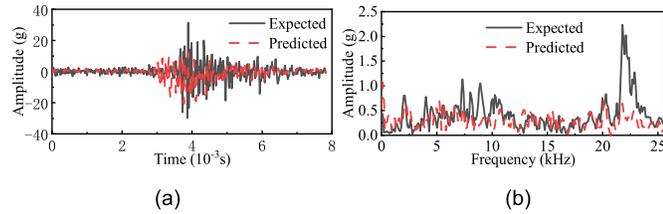


Fig. 6. Vibration signal prediction result. (a) Time-domain. (b) Frequency-domain.

The $u_A(t)$ is a complex signal whose real and imaginary components are u_A^r and u_A^i , respectively. The instantaneous amplitude is the magnitude of $u_A(t)$:

$$Amp_{ins} = \sqrt{(u_A^r)^2 + (u_A^i)^2} \quad (30)$$

The instantaneous phase is:

$$Pha_{ins} = \arctan(u_A^i/u_A^r) \quad (31)$$

where arctan is the arctangent function.

The instantaneous frequency is:

$$Fre_{ins} = d(Pha_{ins})/dt \quad (32)$$

The instantaneous amplitude and the instantaneous frequency are predicted by two models separately. Then, the time series can be obtained by:

$$u_A(t) = Amp_{ins} \cdot \cos(Pha_{ins}) \quad (33)$$

where the instantaneous phase is the integral of instantaneous frequency, and the input signal could provide the constant term.

Based on the amplitude-frequency separation, the predicted vibration is shown in Fig. 7, and NRMSEs are listed in Table II. The accuracy of the result is improved significantly, and the high-frequency components remain complete. It shows

TABLE II
NRMSEs OF VIBRATION SIGNAL TESTING RESULTS

Models	Amplitude	Frequency	Average
SVM	1.42	1.31	1.37
MLP	1.23	1.11	1.08
ELM	1.06	1.24	1.14
ENN	1.17	1.22	1.19
TVF-EMD-ENN	0.94	1.17	1.06
LSTM	1.14	1.12	1.13
EEMD-LSTM	0.86	1.11	0.99
GRU	1.29	1.20	1.24
VMD-GRU	0.94	1.17	1.06
Original ESN	1.15	1.01	1.08
IP-ESN	1.11	1.13	1.12
GESN	0.81	1.05	0.93
IESN	0.63	0.51	0.57

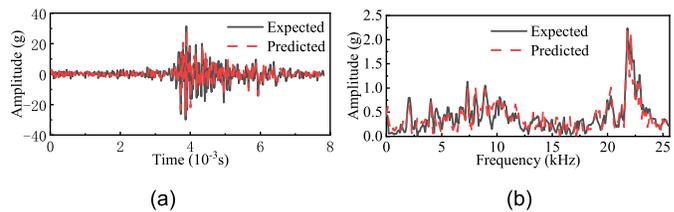


Fig. 7. Vibration signal prediction result by amplitude-frequency separation. (a) Time-domain. (b) Frequency-domain.

the approach can predict high-frequency vibration of the engine to detect the knock.

The testing results of the other models are also listed in Table II. The decomposition algorithm can improve the accuracy in amplitude, but the frequency is still unsatisfactory.

TABLE III
SIGNIFICANCE ANALYSIS BETWEEN THE IESN
AND THE OTHER MODELS

Models	p values	Models	p values
SVM	1.59×10^{-8}	EEMD-LSTM	5.33×10^{-8}
MLP	2.02×10^{-9}	GRU	3.12×10^{-11}
ELM	9.92×10^{-14}	VMD-GRU	8.92×10^{-9}
ENN	3.39×10^{-9}	Original ESN	1.13×10^{-10}
TVF-EMD-ENN	2.89×10^{-16}	IP-ESN	3.05×10^{-8}
LSTM	2.75×10^{-13}	GESN	6.38×10^{-10}

Compared with them, the IESN still obtains the highest accuracy, which means it has a certain prospect in engine health monitoring.

V. ANALYSES AND DISCUSSION

The statistical hypothesis test, echo state property bound, reservoir activation function, reservoir structure, and training data size are discussed further in this section based on the NARMA case shown in Section IV-A.1.

A. Significance Analysis

As mentioned above, the average values of ten runs are taken as the final results of these models; therefore, the t-test is employed to analyze the significance of the IESN result. The p values between the IESN and the other nine models are listed in Table III. The small p values verify the significant difference, which shows the IESN has stable and accurate performance.

B. Bound of the Echo State Property

An IESN used the original echo state property bound, i.e., (7), is tested, in which the hyperparameter is set as $\tilde{\rho} = 0.9$. The prediction result is $\text{NRMSE} = 2.75 \times 10^{-1}$. Compared with it, the complete IESN has a higher accuracy of $\text{NRMSE} = 1.90 \times 10^{-1}$, which shows the superiority of the universal tighter echo state property bound derived in this paper.

C. Reservoir Activation Function

The IESN models with several other reservoir activation functions are tested. The alternative functions are as follows:

The tanh function: $f(x) = (e^x - e^{-x}) / (e^x + e^{-x})$.

The sigmoid function: $f(x) = 1 / (1 + e^{-x})$.

The ReLU function: $f(x) = \max(0, x)$.

The leaky ReLU function: $f(x) = \begin{cases} x & \text{if } x > 0 \\ \lambda x & \text{if } x \leq 0 \end{cases}$, where $\lambda = 0.1$.

The Swish function: $f(x) = x(1 + e^{-\beta x})$, where $\beta = 0.2$.

The Mexican hat wavelet function: $f(x) = (1 - x^2)e^{-x^2/2}$.

Besides, the linear function $f(x) = x$ is also analyzed as the baseline. The Lipschitz constants of these functions are

TABLE IV
TESTING RESULTS OF DIFFERENT ACTIVATION FUNCTIONS

Activation functions	NRMSE
Tanh	3.76×10^{-1}
Sigmoid	4.79×10^{-1}
ReLU	1.09
Leaky ReLU	1.12
Swish	4.07×10^{-1}
Mexican hat wavelet	4.47×10^{-1}
Linear	7.93×10^{-1}
Proposed composite function	1.90×10^{-1}

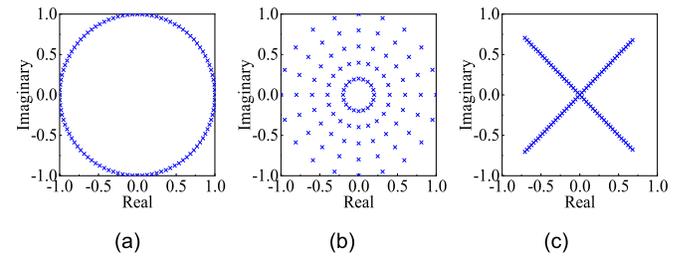


Fig. 8. Eigenvalues distributions in the complex plane. (a) Single circle. (b) Multiple concentric circles. (c) X-shape.

TABLE V
TESTING RESULTS OF DIFFERENT EIGENVALUES DISTRIBUTIONS

Eigenvalues distribution	Single circle	Multiple concentric circles	X-shape	Uniform distribution
NRMSE	3.26×10^{-1}	2.49×10^{-1}	3.12×10^{-1}	1.90×10^{-1}

solved by (16)-(18) to run these models at the edge of chaos. Prediction results are listed in Table IV.

The IESN with the proposed composite function obtains the highest accuracy, which shows the necessity of the research on the reservoir activation function. In addition, the results of the models with the ReLU and leaky ReLU functions are unsatisfactory and even lower than the linear function. It can be deduced that the smooth activation function is beneficial for the IESN.

D. Reservoir Structure

The reservoir should be a normal matrix to obtain the exact maximum structured singular value. Several other possible distribution forms of reservoir eigenvalues in the complex plane are analyzed: the single circle, the multiple concentric circles, and the X-shape, as shown in Fig. 8.

The prediction results of the IESN models with these reservoirs are listed in Table V. The uniform distribution adopted in this paper has the best performance. The other distributions could be ranked in descending order by NRMSEs: the multiple concentric circles, the X-shape, and the single circle. It confirms that varied reservoir eigenval-

TABLE VI
NRMSEs OF TESTING RESULTS BASED ON
DIFFERENT TRAINING DATA SIZE

Training data size	2000	5000	10000
SVM	1.00	9.78×10^{-1}	9.52×10^{-1}
MLP	8.69×10^{-1}	8.65×10^{-1}	8.51×10^{-1}
ELM	1.02	1.01	1.00
ENN	9.72×10^{-1}	9.08×10^{-1}	9.06×10^{-1}
TVF-EMD-ENN	9.23×10^{-1}	9.07×10^{-1}	8.97×10^{-1}
LSTM	1.01	7.74×10^{-1}	7.51×10^{-1}
EEMD-LSTM	8.25×10^{-1}	7.62×10^{-1}	7.53×10^{-1}
GRU	9.99×10^{-1}	8.23×10^{-1}	7.66×10^{-1}
VMD-GRU	8.52×10^{-1}	7.64×10^{-1}	7.42×10^{-1}
Original ESN	8.32×10^{-1}	5.92×10^{-1}	4.94×10^{-1}
IP-ESN	4.13×10^{-1}	3.85×10^{-1}	3.23×10^{-1}
GESN	5.14×10^{-1}	4.22×10^{-1}	4.09×10^{-1}
IESN	1.62×10^{-1}	1.52×10^{-1}	1.39×10^{-1}

ues can obtain abundant dynamic characteristics to improve prediction accuracy and the design of uniform distribution is reasonable.

E. Training Data Size

In the NARMA case, the training and testing datasets contain 1000 samples, respectively. The training data size is adjusted as 2000, 5000, and 10000 to analyze its influence on the prediction accuracy. The NRMSEs of testing results based on different models are listed in Table VI. As the training data size increases, the prediction accuracies of all the models also increase. Nevertheless, the proposed IESN has the best results with different conditions, which shows the model has generality, and its accuracy does not benefit from exceptional testing cases.

VI. CONCLUSION AND OUTLOOK

An IESN is researched in this paper to predict time series accurately and efficiently. A tighter bound of echo state property is deduced firstly for a more precise chaotic edge. A smooth composite activation function is designed to enhance the ESN, and then the exact bound is solved. Finally, a reservoir, whose eigenvalues distribute uniformly in the complex plane, is developed. Compared with several models, the IESN obtains the most accurate results in six benchmarks and a set of vibration signals.

The reservoir is built as a normal matrix for exact structured singular values because of the echo state property bound. Besides, the decoupling places eigenvalues only near the diagonal of the matrix, which results in high sparsity in the large-size matrix. These limit the development of the reservoir, and a more general structure with controllable sparsity will be researched in future work.

REFERENCES

- [1] Z. Wang, H. Qian, D. Zhang, and Y. Wei, "Prediction model design for vibration severity of rotating machine based on sequence-to-sequence neural network," *Math. Problems Eng.*, vol. 2019, Sep. 2019, Art. no. 4670982.
- [2] F. Meng, X. Gong, W. Du, L. Wang, F. Zhao, and L. Li, "Vibration performance prediction and reliability analysis for rolling bearing," *J. Vibroeng.*, vol. 23, no. 2, pp. 327–346, Jan. 2021.
- [3] H. Su, X. Li, B. Yang, and Z. Wen, "Wavelet support vector machine-based prediction model of dam deformation," *Mech. Syst. Signal Process.*, vol. 110, pp. 412–427, Sep. 2018.
- [4] V. Ravi, D. Pradeepkumar, and K. Deb, "Financial time series prediction using hybrids of chaos theory, multi-layer perceptron and multi-objective evolutionary algorithms," *Swarm Evol. Comput.*, vol. 36, pp. 136–149, Oct. 2017.
- [5] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.
- [6] W. Fu, K. Wang, C. Li, and J. Tan, "Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM," *Energy Convers. Manage.*, vol. 187, pp. 356–377, May 2019.
- [7] L. Xu, X. Yu, and T. A. Gulliver, "Intelligent outage probability prediction for mobile IoT networks based on an IGWO-Elman neural network," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1365–1375, Feb. 2021.
- [8] N. Ghadimi, A. Akbarimajd, H. Shayeghi, and O. Abedinia, "Two stage forecast engine with feature selection technique and improved meta-heuristic algorithm for electricity load forecasting," *Energy*, vol. 161, pp. 130–142, Oct. 2018.
- [9] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014, *arXiv:1412.3555*.
- [10] H. Yan, Y. Qin, S. Xiang, Y. Wang, and H. Chen, "Long-term gear life prediction based on ordered neurons LSTM neural networks," *Measurement*, vol. 165, Dec. 2020, Art. no. 108205.
- [11] S. Xiang, Y. Qin, C. Zhu, Y. Wang, and H. Chen, "LSTM networks based on attention ordered neurons for gear remaining life prediction," *ISA Trans.*, vol. 106, pp. 343–354, Nov. 2020.
- [12] W. Xiang, F. Li, J. Wang, and B. Tang, "Quantum weighted gated recurrent unit neural network and its application in performance degradation trend prediction of rotating machinery," *Neurocomputing*, vol. 313, pp. 85–95, Nov. 2018.
- [13] C. Song and X. Chen, "Performance comparison of machine learning models for annual precipitation prediction using different decomposition methods," *Remote Sens.*, vol. 13, no. 5, Mar. 2021, Art. no. 1018.
- [14] Y. Yang and Y. Yang, "Hybrid method for short-term time series forecasting based on EEMD," *IEEE Access*, vol. 8, pp. 61915–61928, 2020.
- [15] C. Li, G. Tang, X. Xue, A. Saeed, and X. Hu, "Short-term wind speed interval prediction based on ensemble GRU model," *IEEE Trans. Sustain. Energy*, vol. 11, no. 3, pp. 1370–1380, Jul. 2020.
- [16] H. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 5667, pp. 78–80, Apr. 2004.
- [17] L. Zhang, C. Hua, Y. Tang, and X. Guan, "Ill-posed echo state network based on l-curve method for prediction of blast furnace gas flow," *Neural Process. Lett.*, vol. 43, no. 1, pp. 97–113, Feb. 2016.
- [18] Y. Yang, X. Zhao, and X. Liu, "A novel echo state network and its application in temperature prediction of exhaust gas from hot blast stove," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 12, pp. 9465–9476, Dec. 2020.
- [19] H.-C. Chen and D.-Q. Wei, "Chaotic time series prediction using echo state network based on selective opposition grey wolf optimizer," *Nonlinear Dyn.*, vol. 104, no. 4, pp. 3925–3935, May 2021.
- [20] X. Sun, T. Li, Q. Li, Y. Huang, and Y. Li, "Deep belief echo-state network and its application to time series prediction," *Knowl.-Based Syst.*, vol. 130, pp. 17–29, Aug. 2017.
- [21] H. Jaeger, M. Lukoševičius, D. Popovici, and U. Siewert, "Optimization and applications of echo state networks with leaky-integrator neurons," *Neural Netw.*, vol. 20, no. 3, pp. 335–352, May 2007.
- [22] X. Wang, Y. Jin, and K. Hao, "Echo state networks regulated by local intrinsic plasticity rules for regression," *Neurocomputing*, vol. 351, pp. 111–122, Jul. 2019.
- [23] J. Qiao, F. Li, H. Han, and W. Li, "Growing echo-state network with multiple subreservoirs," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 2, pp. 391–404, Feb. 2017.
- [24] C. Gallicchio, A. Micheli, and L. Pedrelli, "Deep reservoir computing: A critical experimental analysis," *Neurocomputing*, vol. 268, pp. 87–99, Dec. 2017.

- [25] B. Zhang and Y. Wang, "Echo state networks with decoupled reservoir states," in *Proc. IEEE Workshop Mach. Learn. Signal Process.*, Oct. 2008, pp. 444–449.
- [26] M. Buehner and P. Young, "A tighter bound for the echo state property," *IEEE Trans. Neural Netw.*, vol. 17, no. 3, pp. 820–824, May 2006.
- [27] J. Zhao, C. Zhao, F. Zhang, G. Wu, and H. T. Wang, "Water quality prediction in the waste water treatment process based on ridge regression echo state network," in *Proc. 2nd Int. Conf. Artif. Intell. Appl. Technol.*, vol. 435, 2018, Art. no. 012025.
- [28] G. Zhang, C. Zhang, and W. Zhang, "Evolutionary echo state network for long-term time series prediction: On the edge of chaos," *Appl. Intell.*, vol. 50, no. 3, pp. 893–904, Mar. 2020.
- [29] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. 14th Int. Conf. Artif. Intell. Statist.*, vol. 15, 2011, pp. 315–323.
- [30] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," 2017, *arXiv:1710.05941*.
- [31] G. T. Ribeiro, V. C. Mariani, and L. D. S. Coelho, "Enhanced ensemble structures using wavelet neural networks applied to short-term load forecasting," *Eng. Appl. Artif. Intell.*, vol. 82, pp. 272–281, Jun. 2019.
- [32] F. Bi, X. Li, C. Liu, C. Tian, T. Ma, and X. Yang, "Knock detection based on the optimized variational mode decomposition," *Measurement*, vol. 140, pp. 1–13, Jul. 2019.



Xin Li received the master's degree in power machinery and engineering from Tianjin University, Tianjin, China, in 2017, where he is currently pursuing the Ph.D. degree with the State Key Laboratory of Engines. His main research directions are reliability and weak power fault diagnosis.



Fengrong Bi received the Ph.D. degree in power machinery and engineering from Tianjin University, Tianjin, China, in 2003. He is a Professor with Tianjin University. His main research directions are vehicle dynamics, vibration and noise control, reliability, and weak power fault diagnosis.



Xiao Yang received the degree from Tianjin University, Tianjin, China, where he is currently pursuing the Ph.D. degree in power machine and engineering with the State Key Laboratory of Engines. His current research interests include fault diagnosis and noise reduction of internal combustion engines.



Xiaoyang Bi received the Ph.D. degree in general mechanics and fundamentals of mechanics from Tianjin University, Tianjin, China, in 2019. She is currently a Postdoctoral Researcher with the Hebei University of Technology, Tianjin. Her current research interests include vibration control, fault diagnosis, and deep learning.