Nonlinear Dyn (2023) 111:14295–14308
https://doi.org/10.1007/s11071-023-08592-7

ORIGINAL PAPER

# Noise resistance of next-generation reservoir computing: a comparative study with high-order correlation computation

**Shengyu Liu · Jinghua Xiao · Zixiang Yan · Jian Gao**

Received: 9 February 2023 / Accepted: 11 May 2023 / Published online: 2 June 2023


**Abstract** Reservoir computing (RC) methods have received more and more attention and applications in chaotic time series prediction with their simple structure and training method. Recently, the next-generation reservoir computing (NG-RC) method has been proposed by Gauthier et al. (Nat Commun 12:5564, 2021) with less training cost and better time series predictions. Nevertheless, in practice, available data on dynamic systems are contaminated with noise. Though NG-RC is shown highly efficient in learning and predicting, its noise resistance captivity is not clear yet, limiting its use in practical problems. In this paper, we study the noise resistance of the NG-RC method, taking the well-known denoising method, the high-order correlation computation (HOCC) method, as a reference. Both methods have similar procedures in respect of function bases and regression processes. With the simple ridge regression method, the NG-RC method has a strong noise resistance for white noise, even better than the HOCC method. Besides, the NG-RC method also shows a good prediction ability for small colored noise, while it does not provide correct reconstruct dynamics. In this paper, other than reconstruction parameters, four numerical indicators are used to check the noise resistance comprehensively, such as the training error, prediction error, prediction time, and auto-correlation prediction error, for both the short-time series and long climate predictions. Our results provide a systematic estimation of NG-RC's noise resistance capacity, which is helpful for its applications in practical problems.

**Keywords** Next-generation reservoir computing · Noise resistance · Reservoir computing · High-order correlation computation

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s11071-023-08592-7.

S. Liu · J. Xiao · Z. Yan · J. Gao (✉)
School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China
e-mail: gao.jian@bupt.edu.cn

S. Liu
e-mail: liushengyv@yeah.net

J. Xiao
e-mail: jhxiao@bupt.edu.cn

Z. Yan
e-mail: yanzx@bupt.edu.cn

## 1 Introduction

Analysis and prediction of data play an important role in people's production and life, such as weather prediction, environmental pollution control, earthquake prediction, financial data analysis, speech recognition, image processing, and aircraft control [1–10]. For the field of time series prediction, various

methods have been proposed and obtained more satisfactory prediction results, such as system dynamics reconstruction and neural network prediction. However, in various applications, all the prediction methods are facing a central problem—noise, which brings great difficulties for prediction. For example, when extracting valuable signals from the collected data, the correlation between the noises affects the estimation performance of the signal parameters seriously [11]. In radar monitoring, the presence of noise can make the dynamic reconstruction results unstable [12]. In remote sensing, the effect of noise on the prediction of system dynamics can lead to biased detection of spatiotemporal concentration distribution information [13]. Therefore, the ability to accurately obtain system dynamics from noisy data becomes one of the essential indicators of prediction methods.

To overcome the effects caused by noise in the data, researchers have proposed a variety of dynamical reconstruction methods, such as smoothing method [14], polynomial fitting method [15], local dynamics global fitting method [16], and high-order correlation computation (HOCC) method [17]. The smoothing method takes multi-step averaging to attenuate the noise effect. The polynomial fitting method proposed by Lu et al. [15] directly uses polynomials to fit the time series with noise. The local dynamics global fitting method proposed by Wang and Lan et al. [16] uses globally invariant polynomials to fill all the local pieces of time series. The HOCC method proposed by Chen et al. uses the differential-time correlation of variables to filter out the noise and solves all the unknown coefficients in the system equation by calculating the high-order correlation between variables. In HOCC, the use of differential-time correlation to remove the effect of noise can adjust the time difference in a considerable range to adapt to different noise conditions.

In the last decades, lots of new methods have been proposed to learn, predict and reconstruct the dynamics from data, such as different network structures detection methods [18, 19], extra local driving for topology inference [20], and the compressive sensing technology [21, 22]. Among all these new methods, neural networks have also attracted much attention and applications in the field of prediction due to their good nonlinear mapping capability, self-learning adaptation ability, and parallel information processing capability [23–27]. At the beginning of this century, a novel recurrent neural network method, reservoir computing (RC), was proposed to bring a breakthrough to chaotic forecasting with its simple structure and training method [28–34]. Recently, the next-generation reservoir computing (NG-RC) method, proposed by Gauthier et al. [35] simplifies the RC computation system, significantly reducing its demand on computer resources and saving a lot of time. The new method creates linear and nonlinear feature vectors directly from discretely sampled input data without using a neural network. In NG-RC, the linear feature vector consists of constant terms and observations of the input vector at the current and certain previous time steps. The nonlinear feature vector consists of a two-by-two combination of linear components. The NG-RC method is 33–162 times faster than traditional RC calculations and requires only 28 neurons to achieve the accuracy that would have been achieved with 4000 neurons. Besides, the new method uses 400 data points to obtain the same results as the traditional RC using 5000 or even more data points for training (the exact number of data points depends on the required accuracy).

The high efficiency of the NG-RC methods attracts lots of attention. However, as one of the learning and prediction methods, the noise resistance of NG-RC is not discussed in detail yet, limiting its applications in practice where the available data is always contaminated with noise. Different from classical reservoir computing methods, the NG-RC method depends on the feature vectors, which is similar to the function bases of some reconstruction methods, such as HOCC. As one of the well-developed noise-resistant methods, the effectiveness, robustness, and adaptability of HOCC to different conditions have also been studied comprehensively in simulation validation [17, 36, 37]. Hence, to study the noise resistance of NG-RC comprehensively, we take the HOCC as a reference and compare the differences between these two methods and their noise resistance ability. Surprisingly, we find that even though the NG-RC method does not have special designs for noise resistance, it surpasses the HOCC method for systems with white noise and provides reasonable prediction ability for small colored noise.

In this paper, we compare the noise resistance ability and characteristics of NG-RC with reference to HOCC methods from theoretical analysis and numerical experiments. In terms of theory, we analyze the

similarities between the two methods. Both methods have similar procedures in the respect of function bases and regression processes, especially when one considers the NG-RC method without time-delay function bases. We take the Lorenz system as an example and explore the difference in the coefficient reconstructed by the two methods. In terms of numerical experiments, four indicators are introduced to show the noise resistance of methods, such as the training error, prediction error, prediction time, and auto-correlation prediction error. Both the white and colored noise are considered in this paper. The effects of the noise intensity, training length, and sampling interval on the noise resistance ability of the two methods are explored comprehensively. Besides, the noise resistance to colored noise with the change of time delay in NG-RC and HOCC is also studied to study the potential variations of the method for colored noise.

This paper is organized as follows: Section 2 introduces NG-RC and HOCC methods and compares their difference theoretically. Sections 3 and 4 show the numerical experimental results of the NG-RC method without and with time-delay function bases on the white and colored noise-driven system. Section 5 concludes the paper.

## 2 Theoretical comparative analysis of NG-RC and HOCC

Considering the dynamics of an arbitrary noise-driven system as:

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{f}(\boldsymbol{x}(t)) + \boldsymbol{\Gamma}(t), \\ \boldsymbol{x} &= (x_1, x_2, \ldots, x_M)^T, \\ \boldsymbol{f} &= (f_1, f_2, \ldots, f_M)^T, \\ \boldsymbol{\Gamma} &= (\Gamma_1, \Gamma_2, \ldots, \Gamma_M)^T, \end{aligned} \tag{1}$$

where $\boldsymbol{x}$ is the state variable, $M$ the dimensionality of the state variable, $t$ the time variable, and $\boldsymbol{f}$ the function. The $\Gamma_i(t), i = 1, 2, \ldots, M$ is the noise term which is either Gaussian white with $<\boldsymbol{\Gamma}(t)> = 0$ and $<\boldsymbol{\Gamma}(t)\boldsymbol{\Gamma}(s)> = 2D\delta(t-s)$ where $D$ is the noise intensity, or Ornstein–Uhlenbeck colored noise with $<\boldsymbol{\Gamma}(t)\boldsymbol{\Gamma}(s)> = \frac{1}{2\tau}\exp(-|t-s|/\tau)$ in this paper. From the dynamical system Eq. (1), one gets the measurable data as time series,

$$\begin{aligned} &\boldsymbol{x}(t_1), \boldsymbol{x}(t_2), \ldots, \boldsymbol{x}(t_k), \ldots, \boldsymbol{x}(t_N), \\ &\Delta t = t_{k+1} - t_k \ll 1; k = 1, 2, \ldots, N-1; N \gg 1, \end{aligned} \tag{2}$$

where $N$ denotes the time series length, and $\Delta t$ is the time sampling interval.

In this paper, we consider the inverse problem that given the time series in Eq. (2), how to obtain the original dynamical system Eq. (1). Here, we take the Lorenz system [5] as the example of an original noise-driven dynamic system. It is one of the most famous models in chaos studies, developed in 1963 as the first system discovered to produce chaotic attractors. The Lorenz system follows the dynamics of Eq. (1) with three coupled nonlinear differential equations:

$$\begin{aligned} \boldsymbol{f} &= (f_1, f_2, f_3)^T \\ &= (\sigma(y-x), x(\rho - z) - y, xy - \beta z)^T, \end{aligned} \tag{3}$$

where $x$, $y$, $z$ denote the state variables. In this paper, we take the parameters $\sigma = 10, \rho = 28, \beta = 8/3$ to for chaotic time series. In the following, we use NG-RC and HOCC methods for the inverse problems with noise to get the original Lorenz dynamical systems.

### 2.1 NG-RC method

The NG-RC method is developed from the traditional RC. It is no longer requiring a linear combination of the input signals using a randomly generated neural network to obtain the output signal, but instead directly creates feature vectors with discretely sampled input data, where the feature vectors are called bases. For prediction, the bases consist of three parts: constant term, linear terms, and nonlinear terms. For the Lorenz system, the linear terms at the moment $t_k$ are usually composed of $x(t_k), y(t_k), z(t_k), x(t_{k-1}), y(t_{k-1}), z(t_{k-1})$ [35], where $k = 2, \ldots, N-1$. The nonlinear terms component is composed of two combinations of constant and linear terms, totaling 21 terms. The whole basis function matrix $\boldsymbol{P}$ is obtained from the observed data $\boldsymbol{x}$, where the $k$-th column $\boldsymbol{P_k}$ for time $t_k$ is:

$$\begin{aligned}\boldsymbol{P_k} = (&1, x(t_k), y(t_k), z(t_k), x(t_{k-1}), y(t_{k-1}),\\ &z(t_{k-1}), x^2(t_k), x(t_k)y(t_k), x(t_k)z(t_k),\\ &x(t_k)x(t_{k-1}), x(t_k)y(t_{k-1}), x(t_k)z(t_{k-1}),\\ &y(t_k)y(t_{k-1}), y(t_k)z(t_{k-1}), z^2(t_k), y^2(t_k),\\ &y(t_k)z(t_k), y(t_k)x(t_{k-1}), z(t_k)x(t_{k-1}),\\ &z(t_k)y(t_{k-1}), z(t_k)z(t_{k-1}), x^2(t_{k-1}),\\ &x(t_{k-1})y(t_{k-1}), x(t_{k-1})z(t_{k-1}),\\ &y^2(t_{k-1}), y(t_{k-1})z(t_{k-1}), z^2(t_{k-1})).\end{aligned} \tag{4}$$

The basic assumption of NG-RC is that the function base is complete for the original dynamical systems, from which the dynamical function $\boldsymbol{f}(\boldsymbol{x})$ in Eq. (1) could be expressed as a linear transformation of $\boldsymbol{P}$ as

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A_1 P}, \tag{5}$$

where $\boldsymbol{A}_1$ is the coefficient matrix. The left hand side of Eq.(5) can be estimated through a simple difference method as:

$$\boldsymbol{f}(\boldsymbol{x_k}) \approx \boldsymbol{X_k} = \frac{1}{\Delta t}\begin{pmatrix} x(t_{k+1}) - x(t_k) \\ y(t_{k+1}) - y(t_k) \\ z(t_{k+1}) - z(t_k) \end{pmatrix}. \tag{6}$$

Taking $\boldsymbol{X_k}$ as the $k$-th column, one gets the target matrix $\boldsymbol{X}$ whose size is $3 \times L$, where $L = N - 2$ is the length of training data. Correspondingly, the size of coefficient the matrix $\boldsymbol{A}_1$ and basis function matrix $\boldsymbol{P}$ are $3 \times 28$ and $28 \times L$ respectively.

In NG-RC method, the ridge regression method is used to solve $\boldsymbol{A}_1$:

$$\boldsymbol{A_1} = \boldsymbol{X P^T}(\boldsymbol{P P^T} + \alpha \boldsymbol{I})^{-1}, \tag{7}$$

where $\boldsymbol{P^T}$ is matrix transpose of $\boldsymbol{P}$, $\alpha$ ridge regression parameter, and $\boldsymbol{I}$ the identity matrix. After solving the reserve problem and getting $\boldsymbol{A}_1$, an Euler-like integration step can be used to obtain the prediction time series $\boldsymbol{y}$:

$$\boldsymbol{y_{i+1}} = \boldsymbol{y_i} + \boldsymbol{A_1 P_i} \Delta t, \tag{8}$$

with $\boldsymbol{i} = 1, 2, \ldots, \infty$. For learning and prediction of chaotic time series, the NG-RC method using bases with time delay is shown to have the same or even better prediction capacity compared to the traditional RC using randomly generated neural networks [35]. But here we focus on its noise resistance ability.

## 2.2 HOCC method

In this paper, we take the HOCC method as the reference, which uses the differential-time correlations of variables to remove noise and solves all unknown coefficients in the system equation by calculating the high-order correlations among the variables. When using the HOCC method for prediction, it is necessary to select an appropriate function basis in advance, denoted as $\boldsymbol{Q}$ as not to be confused with the function basis for NG-RC. For the Lorenz system, the polynomial functions are often chosen as the bases. Similar to the NG-RC method, the whole basis function matrix $\boldsymbol{Q}$ is obtained from the observed data $\boldsymbol{x}$, where the $k$-th column $\boldsymbol{Q_k}$ for time $t_k$ as:

$$\begin{aligned}\boldsymbol{Q_k} = (&1, x(t_k), y(t_k), z(t_k), x^2(t_k), x(t_k)y(t_k),\\ &x(t_k)z(t_k), y^2(t_k), y(t_k)z(t_k), z^2(t_k)).\end{aligned} \tag{9}$$

It is worth to be noted that the basis function matrix $\boldsymbol{Q}$ is exactly a part of the function matrix $\boldsymbol{P}$ except for the time-delay terms.

The basic assumption of HOCC is also the same as NG-RC, where the function base $\boldsymbol{Q}$ is complete for the original dynamical systems, from which the dynamical function $\boldsymbol{f}(\boldsymbol{x})$ in Eq. (1) could be expressed as a linear transformation of $\boldsymbol{Q}$ as

$$\boldsymbol{f}(\boldsymbol{x}(t)) = \boldsymbol{A_2 Q}, \tag{10}$$

where $\boldsymbol{A}_2$ denotes the coefficient matrix of $\boldsymbol{Q}$. Up to this step, the two methods, NG-RC and HOCC, are exactly the same, with only a slightly difference in the choice of function basis. If we use the ridge regression method to get $\boldsymbol{A_2}$ directly, this is the NG-RC method without time-delay terms. However, in the HOCC method, one more step is applied to remove the noise. Here, one takes the function vector $\boldsymbol{Q}^T(\boldsymbol{x}(t-\tau))$ with time delay as

$$\boldsymbol{Q}(\boldsymbol{x}) = (Q_1(\boldsymbol{x}), Q_2(\boldsymbol{x}), \ldots, Q_{10}(\boldsymbol{x}))^T, \tag{11}$$

and right multiple it with Eq. (1). Then by taking time averaging to calculate all related correlations, we have

$$\boldsymbol{B}^T(-\tau) = \boldsymbol{A_2}\hat{\boldsymbol{C}}(-\tau) + <\boldsymbol{\Gamma}(t)\boldsymbol{Q}^T(t-\tau)>, \tag{12}$$

with

$$
\begin{aligned}
\boldsymbol{B}(-\tau) &= (B_1(-\tau), B_2(-\tau), \ldots, B_{10}(-\tau))^T, \\
B_\mu(-\tau) &= <\dot{x}(t) Q_\mu(t-\tau)> \\
&= \frac{1}{L-P} \sum_{k=p+1}^{L} \dot{x}(t_k) Q_\mu(t_{k-p}), \\
\dot{x}(t_k) &= \frac{x(t_{k+1}) - x(t_k)}{\Delta t}, \tau = p\Delta t, \\
\hat{\boldsymbol{C}}(-\tau) &= (C_{\mu\upsilon}) \\
&= \left( \frac{1}{L-P} \sum_{k=p+1}^{L} Q_\mu(t_k) Q_\upsilon(t_{k-p}) \right),
\end{aligned}
\tag{13}
$$

where $\mu = 1, 2, \ldots, 10$, $< \cdot >$ denotes time averaging. The time delay $\tau$ is assumed to satisfy the inequality $0 \approx \tau_d \ll \tau \ll 1$, that it is much larger than the correlation time of dynamical noise $\tau_d$, and much smaller than the characteristic times of deterministic network dynamics, previously assumed to be of order 1. In this case, noises and correlations are decorrelated as

$$
<\boldsymbol{\Gamma}(t) \boldsymbol{Q}^T(t-\tau)> \approx 0, \tau > t_d, \tag{14}
$$

since the fast-changing noise must not be correlated with any variable data of previous times, disregarding any forms of colored noises. Now with the noise-decorrelation of Eqs. (14) and (12) can be reduced to

$$
\boldsymbol{B}^T(-\tau) = \boldsymbol{A_2} \hat{\boldsymbol{C}}, \tag{15}
$$

which leads to

$$
\boldsymbol{A_2} = \boldsymbol{B}^T(-\tau) \hat{\boldsymbol{C}}^{-1}, \tag{16}
$$

which could be solved directly. This differential-time correlation-based procedure to get $\boldsymbol{A}_2$ is the essential part of the HOCC method and also the major difference between the NG-RC method. After getting the coefficient matrix $\boldsymbol{A}_2$, one can reconstruct the original dynamic is as

$$
\dot{\boldsymbol{x}}(t) = \boldsymbol{A_2} \boldsymbol{Q}, \tag{17}
$$

and also iterate the system and obtain the predicted time series $\boldsymbol{y}$ similar to NG-RC.

It can be seen that both the NG-RC and HOCC methods use the same idea to fit the original system with function bases that are selected in advance and then solve its coefficient matrix. If both methods choose the same bases, the target matrix for its optimization is the same. The difference is that in the HOCC method, the $\hat{\boldsymbol{C}}$ in Eq. (15) is composed of the differential-time correlations of the basis vectors, and the coefficient matrix $\boldsymbol{A}_2$ is regression solved after the correlations, so that the noise is removed by using the property that the noise and the correlators must be decorrelated. While in the NG-RC method, the matrix $\boldsymbol{Y}$ is directly combined by the basis vectors. The ridge regression method is used directly to solve the coefficient matrix $\boldsymbol{A}_1$, and by adjusting $\alpha$ to remove noises while preventing over-fitting.

It is straightforward to notice that these two methods are comparable in theory. In the following, we will use numerical methods to check how well the simple ridge regression method used in NG-RC is for the noise resistance, compared to the HOCC method which has a special designed time difference correlations procedure for noise elimination.

## 3 The NG-RC method without time-delay function bases

In this paper, the Lorenz system in Eq. (3) is set with $\sigma = 10$, $\rho = 28$, $\beta = 8/3$(the Lorenz system is chaotic with this parameter), and the sampling interval reads $h = 0.0002$. The data of the original Lorenz system is generated by the iteration of the Runge–Kutta method, which is noted as $\boldsymbol{y}_{\text{Lorenz}}$. Two types of noise are considered, as the Gaussian white noise and Ornstein–Uhlenbeck colored noise. To show the noise effect, we increase the noise strength gradually, from 0.0001 to 25. The colored noise correlation time is fixed at 25 h. The phase diagrams with either white or colored noise are shown in Fig. 1.

It is clear that noises affect the dynamics of the Lorenz system dramatically in Fig. 1. How to get the original system through such data is the challenging problem that we want to solve with NG-RC and HOCC methods.

In Section 2, we know that if the same function bases are selected, i.e., the NG-RC method without time-delay function bases, then the goal matrix of the two methods is the same. But NG-RC and HOCC have different procedures for noise resistance. Hence, in the following, we compare NG-RC and HOCC in two steps. In this section, we only focus on the NG-RC method without time-delay function bases, denoted as

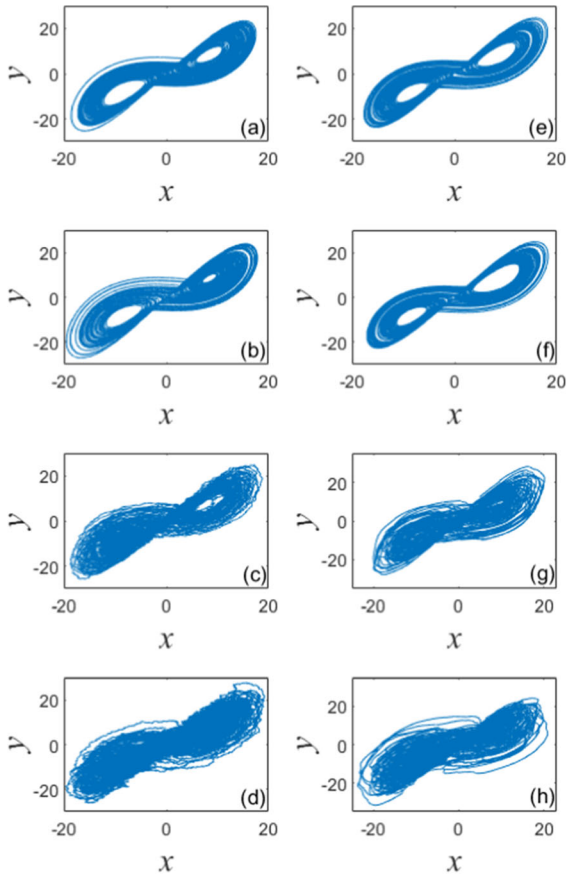


**Fig. 1** Phase diagrams of the Lorenz system are driven by different intensities of white noise or colored noise. $h = 0.00002$, $\sigma = 10$, $\rho = 28$, $\beta = 8/3$. (**a**–**d**) show the Lorenz system driven by white noise with $D = 0.0001$, $D = 0.01$, $D = 5$, and $D = 15$. (**e**–**h**) show the Lorenz system driven by colored noise with $D = 0.0001$, $D = 0.01$, $D = 5$, and $D = 15$

NG-RC10 for it only considering the 10 function bases. And in the next section, we will use the full bases introduced in NG-RC.

### 3.1 White noise-driven system

First, consider the white noise-driven Lorenz system. With HOCC and NG-RC methods, the dynamical system is reconstructed. Compared with the original Lorenz system, We obtain reconstruction errors of these two methods and show them in Table 1.

When the $D$ is smaller, the standard deviation of the coefficients corresponding to the Lorenz system obtained by the HOCC method is about $10^{-5}$ to $10^{-6}$, and the other terms are about $10^{-2}$. When the $D$ is large, the standard deviation of coefficient terms is about $10^{-3}$, and other terms are about $10^{-2}$. It can be found that the reconstruction result of NG-RC10 is close to that of the HOCC method. HOCC method has only slightly better noise resistance ability than NG-RC10 for white noise cases. Here, the train steps $L$ and sampling interval $h$ are chosen as $L = 15$ million and $h = 0.0002$, which are chosen as one of the best parameters for the HOCC method, studied in [37].

To further explore the processing ability of the two methods for noisy data with different parameters, and better understand their noise resistance characteristics, we adopted the control variable method to analyze the noise resistance performance of the NG-RC and HOCC methods from the perspective of the influence of train steps $L$, sampling interval $h$ and noise strength $D$ on the predicted results respectively. Instead of the reconstructed parameters, here we use a single numerical indicator introduced in [35], $E_t$, to estimate the reconstruction result, which reads

$$E_t = \frac{1}{N}\sum_{k=1}^{N}(y_{\text{Lorenz}_k} - y_k)^2, \tag{18}$$

as mean square error between reconstructed time series $\boldsymbol{y}$ and the original one $\boldsymbol{y}_{\text{Lorenz}}$. The smaller the $E_t$ is, the better the noise resistance ability of the method. The experimental results of $E_t$ changes with the above three variables are shown in Fig. 2. Here, considering that the ridge regression parameter $\alpha$ of the NG-RC method has a certain influence on its noise resistance ability, for each data point in Fig. 2, $\alpha$ was optimized to obtain the optimal $E_t$ for display. The optimization process is in Supplementary Note 1.

When the sampling interval is sufficiently small, as shown in Fig. 2a, fixing $h = 0.00002$ and choosing four different intensities of white noise to drive the system for training, the $E_t$ of the two methods do not differ much when the amount of training data is small, and both decrease as the number of train steps increases. These results coincide with the one obtained from the reconstructed parameters shown in Table 1 that the NG-RC10 method and HOCC method have almost the same noise resistance in these cases.

However, as shown in Fig. 2b, increasing the sampling interval to $h = 0.001$, the reconstruction error $E_t$ of the HOCC method appears to be a "plateau area" when $D$ is small, while $E_t$ continues decreasing with the increase of $L$ for NG-RC10 method. The

**Table 1** When the white noise-driven system is reconstructed, the standard deviation of the coefficients obtained by the two methods and the mean of the difference correspond to the original Lorenz system

| | | $D = 0.0001$ | | $D = 0.1$ | | $D = 1.2$ | |
|---|---|---|---|---|---|---|---|
| | | NG-RC10 | HOCC | NG-RC10 | HOCC | NG-RC10 | HOCC |
| $\dot{x}$ | $x$ | $1e-5 \pm 9e-5$ | $-0.03181 \pm 9e-5$ | $-0.014 \pm 3e\text{-}3$ | $-0.03811 \pm 9e-5$ | $0.097 \pm 4e\text{-}3$ | $-0.03811 \pm 9e-5$ |
| | $y$ | $4e-5 \pm 6e-5$ | $0.0111 \pm 6e-4$ | $0.011 \pm 2e-3$ | $0.01112 \pm 6e-5$ | $-0.067 \pm 3e-3$ | $0.01112 \pm 6e-5$ |
| $\dot{y}$ | $x$ | $-6e-5 \pm 8e-5$ | $0.0249 \pm 1e-4$ | $1e-3 \pm 1e-3$ | $0.024 \pm 2e-3$ | $0.050 \pm 5e-3$ | $0.064 \pm 5e-3$ |
| | $y$ | $1.2e-4 \pm 6e-5$ | $-0.03060 \pm 8e-5$ | $-4.2e-3 \pm 9e-4$ | $-0.0337 \pm 9e-4$ | $-0.028 \pm 4e-3$ | $-0.053 \pm 4e-3$ |
| | $xz$ | $0 \pm 2e-5$ | $-1.256e-3 \pm 3e-6$ | $-7e-5 \pm 3e-5$ | $-0.00128 \pm 3e-5$ | $-0.0016 \pm 2e-4$ | $-0.002 \pm 2e-3$ |
| $\dot{z}$ | $z$ | $2.9e-4 \pm 6e-5$ | $-0.01495 \pm 5e-5$ | $5e-3 \pm 1e-3$ | $-0.010 \pm 1e-3$ | $0.008 \pm 3e-3$ | $-0.009 \pm 3e-3$ |
| | $xy$ | $-8e-5 \pm 5e-5$ | $3.076e-3 \pm 4e-6$ | $5e-4 \pm 2e-4$ | $0.0036 \pm 2e-4$ | $-0.004 \pm 1e-3$ | $-0.001 \pm 1e-3$ |
| other bases | | $0 \pm 1e-3$ | $0.01 \pm 0.02$ | $-0.01 \pm 0.02$ | $0.001 \pm 0.03$ | $0 \pm 0.1$ | $0.02 \pm 0.06$ |

simpler NG-RC method works even better than HOCC for time series with large sampling intervals and sufficient large train steps. More details can be found in Fig. 2c, where the training data are the same and large enough (15 million), and the white noise-driven system with three different sampling intervals is selected for training. The reconstruction error $E_t$ of both methods increases with the increase of $D$, while $E_t$ of the HOCC method appears "plateau area" at $h = 0.001$ and $h = 0.0002$. Combined with Fig. 2b, the HOCC method can no longer decrease $E_t$ when training and predicting data with minor $D$ and greater $h$, no matter how to reduce $D$ or increase the number of train steps, and this lower limit varies with $h$. The larger the $h$, the higher the lower limit.

The above results are corroborated in Fig. 2d. For NG-RC10, $E_t$ is more affected by $D$, but the variation of $E_t$ with $h$ is smaller for the same $D$ system. For the HOCC method, when $h$ is small, the smaller $D$ is, the smaller $E_t$ is, similar to the NG-RC10 method. But for large sample interval $h$, the difference between $E_t$ corresponding to different $D$ decreases until it overlaps when $h > 0.001$. In these cases, a decrease in the noise strength $D$ does not result in the improvement of prediction capacity. The sampling interval defines a minimum value of $E_t$, increasing with the increase of $h$. As a result, for time series with large sampling intervals, the NG-RC10 method works better than HOCC, as shown in Fig. 3. Here, the $\alpha$ is also optimized. When $h = 0.00002$, the noise resistance ability of the two methods is comparable in Fig. 3a. When increasing $h$ to 0.0002 and 0.001, the $E_t$ of the HOCC methods becomes larger for smaller intensity noise, while the results of NG-RC10 are less affected, as shown in Fig. 3b, c. The appearance of the "plateau area" of the HOCC method is an open question and out of the scope of this paper, but from this phenomenon, the NG-RC10 method could provide a better prediction for the dynamical system with white noise.

In the above results, the prediction error $E_t$ is used to describe the noise resistance ability of the method. Meanwhile, we noted that other indicators were also used in previous studies to measure the reconstruction or prediction ability of the method. For example, in [16, 17], the results obtained by the reconstruction method are compared with the coefficients of the original dynamic system to measure its reconstruction ability. In [35], the mean square error of the prediction time sequence obtained by the noise resistance method

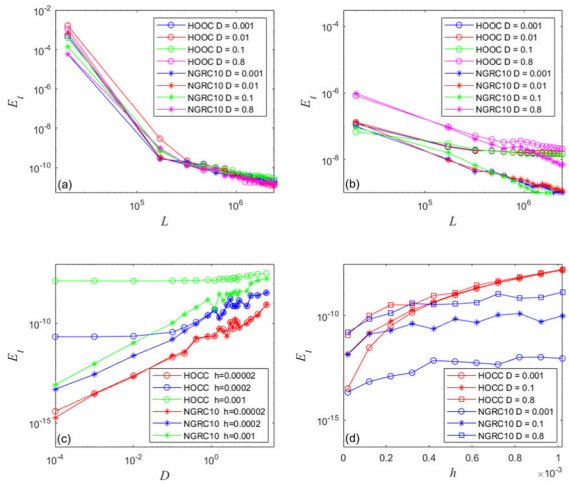


**Fig. 2** Effects of $L$, $D$, and $h$ on the $E_t$ of the two methods. (**a**–**c**) the circle line represents the HOCC method and the star line represents NG-RC10. (**a**, **b**) the influence of $L$ on two methods: blue, red, green, and rose red represents $D$ = 0.001, $D = 0.01$, $D = 0.1$, and $D = 0.8$ respectively, where $h = 0.0002$ in (**a**) and $h = 0.001$ in (**b**). **c** the effect of $D$ on two methods: fixed $L = 15$ million, red, blue and green represents $h = 0.00002$, $h = 0.0002$, $h = 0.001$. **d** the effect of $h$ on two methods: fixed number of $L = 15$ million, the red and blue represent the HOCC and NG-RC10, and the circle, star, and square lines represent the noise intensity of $D = 0.001$, $D = 0.1$, and $D = 0.8$. (Color figure online)

and the training data was calculated to measure the method's fitting ability to the training data. In [38], the prediction ability of the method was measured by the coincident steps between the prediction sequence and the original noise-free sequence. To understand the two methods in multiple dimensions, we check the noise resistance ability of the two methods in terms of a more refined four-dimensional representation. Other than the prediction error $E_t$, we consider the training error $E_n$, prediction time $T_p$, and auto-correlation prediction error $E_a$.

(1) The training error $E_n$, the mean square error of $\boldsymbol{y}$ and $\boldsymbol{y}_{train}$, is defined as

$$E_n = \frac{1}{N}\sum_{k=1}^{N}(y_{train_k} - y_k)^2. \tag{19}$$

It is the objective function of the training phase and is used to measure the learning error with noise.

(2) The prediction time $T_p$, the coincident steps of $\boldsymbol{y}$ and $\boldsymbol{y}_{\text{Lorenz}}$, is defined as the maximum time step $k$ where the square error between the predicted time series and original Lorenz system is smaller than 0.01. It is important to note here that $T_p$ is correlated with $E_t$. A Smaller prediction error $E_t$ results in a larger prediction time $T_p$. But for different dynamical systems, $T_p$ also depends on the system's maximum Lyapunov exponent, where the Lyapunov time is usually used as the unit of $T_p$.

(3) Auto-correlation prediction error $E_a$ is defined as the time series auto-correlation coefficient prediction error between $\boldsymbol{y}$ and $\boldsymbol{y}_{\text{Lorenz}}$. To measure the difference between the predicted time series and the original Lorenz system time series long-time evolution behavior of the two methods. The auto-correlation coefficients of the time series $r_k(k = 1, 2, \ldots, N)$ are calculated using the following equation: $r_k = \frac{c_k}{c_0}$ with $c_k = \frac{1}{T}\sum_{t=1}^{T-k}(x_t - \bar{x})(x_{t+k} - \bar{x})$ and $c_0 = \frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^2$. Here $\{x_t\}$ represents a set of time series, $T$ represents the length of the time series, and $\{x_{t+k}\}$ represents the time series after delaying $\{x_t\}$ by $k$ steps, leading to

$$E_a = \frac{1}{N}\sum_{k=1}^{N}(r_{\text{Lorenz}_k} - r_k)^2. \tag{20}$$

The training and prediction results of the two methods using the above four indexes for the white noise-driven system are shown in Fig. 4, where the two methods use the same data set for training, and the size of the training data is selected to be 15 million.

It can be seen that the two methods have very different performances in the four indicators. For training error $E_n$, there two methods have similar performance, satisfying $E_n = 10^{-5}$ approximately as shown in Fig. 4a, showing that both methods has reached the optimized result as they designed to get. However, such optimized results from these two methods are different, shown as by the prediction error $E_t$ in Fig. 4b, especially when $D$ is small. Here $E_t$ of NG-RC10 is smaller than that of the HOCC method, the dynamical system from NG-RC10 is closer to the original Lorenz system than HOCC. This fact results in a longer prediction time $T_p$ and smaller auto-correlation prediction error $E_a$ of NG-RC10 in Fig. 4c, d, describing the better prediction ability for the short and long evolution of the system.

Compared with the HOCC method, we can conclude that the NG-RC method with the same function bases has a better noise resistance and prediction ability for white noise. This is from the learning limit

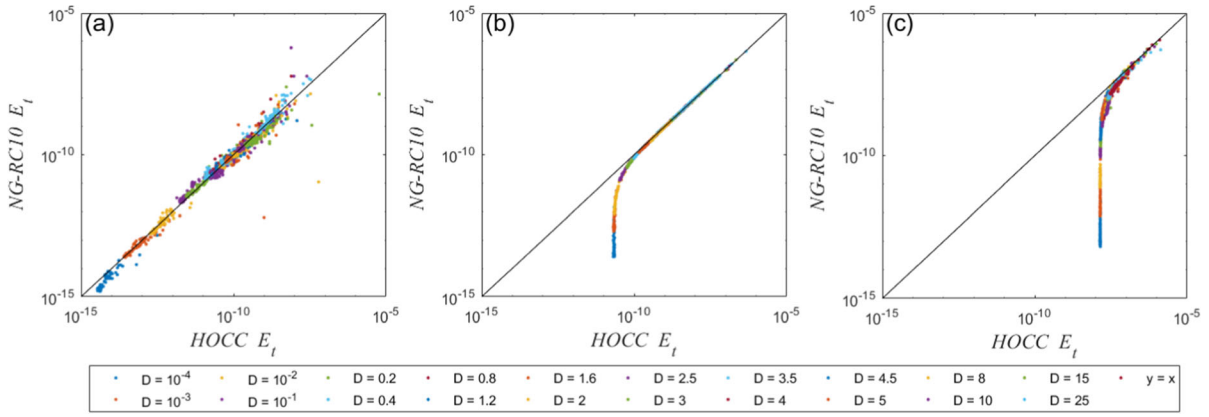


**Fig. 3** Prediction error $E_t$ of the two methods for different noise strength $D$ with sampling interval $h$. The black solid line indicates that the $E_t$ of the two methods are equal. The $h$ used in (**a**–**c**) are $h = 0.001$, $h = 0.0002$ and $h = 0.00002$, $L =$ 15 million, and the dots of different colors represent 20 different $D$. (Color figure online)

of HOCC by large sampling intervals. Studying the mechanism of HOCC is out of the scope of this paper. But for the NG-RC method, it is sufficient to conclude a positive statement for its noise resistance to white noise. We will consider the NG-RC method with time-delay terms in the next section and obtain the same results.

### 3.2 Colored noise-driven system

The spectrum of noise is an important feature. The HOCC method has developed a full frame to deal with colored noise and shows a good noise resistance ability [37]. Here we take the Ornstein–Uhlenbeck noise and check the noise resistance of the NG-RC10 method with colored noise.

Similar to the analysis of white noise, we first reconstruct the parameters and check the reconstruction errors of HOCC and NG-RC10. The results are shown in Table 2. No matter how small the noise strength $D$ is, the reconstruction errors of NG-RC10 are much larger than HOCC. As the variations of the reconstructed parameters are kept sufficiently small, one can conclude that the reconstructed parameters by NG-RC10 have systematic bias from the original dynamical systems. This error is removed by HOCC through differential-time correlation methods.

The training and prediction results of the two methods using the four numerical indicators for the colored noise-driven system are shown in Fig. 5, where the two methods use the same data set for training, and the size of the training data is selected to

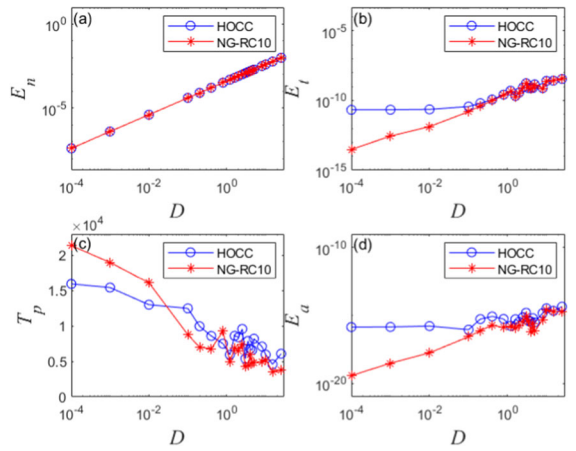


**Fig. 4** $E_n$, $E_t$, $T_p$, $E_a$ of the white noise-driven system by NG-RC10 and HOCC methods. The blue line represents the HOCC method, and the red line represents NG-RC10. (Color figure online)

be 15 million. The optimization process of $\alpha$ is shown in Supplementary Note 2. Similar to the white noise cases, $E_n$ of the HOCC method and NG-RC10 are almost the same, but for the prediction error, $E_t$ of the NG-RC10 method is always larger than $E_t$ of HOCC. Consequently, HOCC has better prediction ability when dealing with the prediction time $T_p$ and auto-correlation prediction error $E_a$. However, it is interesting to note that when the noise strength is sufficiently small as $D < 0.01$, the NG-RC10 and HOCC have almost the same numerical indicators, showing the same prediction capacity, even the reconstructed parameters shown in Table 2 from

**Table 2** When the colored noise-driven system is reconstructed, the standard deviation of the coefficients obtained by the two methods and the mean of the difference correspond to the original Lorenz system

| | | $D = 0.0001$ | | $D = 0.1$ | | $D = 1.2$ | |
|---|---|---|---|---|---|---|---|
| | | NG-RC10 | HOCC | NG-RC10 | HOCC | NG-RC10 | HOCC |
| $\dot{x}$ | $x$ | $0.0385 \pm 2e\text{-}4$ | $5e-4 \pm 2e\text{-}4$ | $0.188 \pm 3e\text{-}3$ | $5e-4 \pm 2e\text{-}4$ | $1.12 \pm 0.02$ | $5e-4 \pm 2e\text{-}4$ |
| | $y$ | $-0.0113 \pm 1e\text{-}4$ | $-3e-4 \pm 1e\text{-}4$ | $-0.090 \pm 2e\text{-}3$ | $-2e-4 \pm 1e\text{-}4$ | $-0.61 \pm 0.01$ | $-3e-4 \pm 1e\text{-}4$ |
| $\dot{y}$ | $x$ | $-0.0242 \pm 3e\text{-}4$ | $5e-4 \pm 2e\text{-}4$ | $0.131 \pm 3e\text{-}3$ | $0.148 \pm 3e\text{-}3$ | $1.07 \pm 0.02$ | $1.07 \pm 0.02$ |
| | $y$ | $0.03027 \pm 1e\text{-}5$ | $-3e-4 \pm 1e\text{-}4$ | $-0.052 \pm 2e\text{-}3$ | $-0.078 \pm 2e\text{-}3$ | $-0.58 \pm 0.01$ | $-0.59 \pm 0.01$ |
| | $xz$ | $1.239e-3 \pm 8e\text{-}6$ | $-1.5e-5 \pm 5e\text{-}6$ | $-2.85e-3 \pm 9e\text{-}5$ | $-3.0e-3 \pm 8e\text{-}5$ | $-0.0288 \pm 7e\text{-}4$ | $-0.0292 \pm 7e\text{-}4$ |
| $\dot{z}$ | $z$ | $0.01674 \pm 8e\text{-}5$ | $1.62e-3 \pm 8e\text{-}5$ | $-0.146 \pm 4e\text{-}3$ | $-0.161 \pm 4e\text{-}3$ | $-1.371 \pm 9e\text{-}3$ | $-1.38 \pm 9e\text{-}3$ |
| | $xy$ | $-3.35e-3 \pm 2e\text{-}5$ | $-1.8e-4 \pm 1e\text{-}5$ | $0.02165 \pm 6e\text{-}5$ | $0.0246 \pm 6e\text{-}4$ | $0.194 \pm 1e\text{-}3$ | $0.195 \pm 1e\text{-}3$ |
| Other bases | | $-0.01 \pm 3e\text{-}2$ | $-2e-3 \pm 5e\text{-}3$ | $0.1 \pm 0.4$ | $0.1 \pm 0.3$ | $1 \pm 3$ | $1 \pm 3$ |

NG-RC10 have a much larger error compared with the ones of HOCC.

The power of HOCC to deal with colored noise is from the time delay $\tau$ in the differential-time correlations. When $\tau$ is much larger than the correlation length $\tau_d$ of the noise and much smaller than the characteristic time of the system dynamics, HOCC has a good noise resistance of colored noise [4, 17]. As shown in Fig. 6, the prediction error $E_t$ decreases with the increase of $\tau$ to a minimum value after $\tau > 100$. As for the NG-RC10 method, its prediction error $E_t$ is the same as the HOCC with $\tau = 1$.

The NG-RC10 methods without the differential-time correlations could not reconstruct the original dynamics correctly, as it mixes the dynamics and noise correlations. However, if the noise strength is sufficiently small, the NG-RC10 method could give the same good prediction of the system's evolution as the ones from the HOCC method, which has good reconstructed parameters. The mechanism of this phenomenon is an open question, and out of the scope of this paper. We will study it in our following work with various types of colored noise.

## 4 The NG-RC method with time-delay function bases

In the above section, NG-RC10 shows a good effect in antiwhite noise prediction and even performs better than the HOCC method when the sampling interval $h$ is large. However, in colored noise resistance prediction, its noise resistance ability is worse than HOCC, especially when the noise strength $D$ is large. In this section, we further consider the NG-RC method with all 28 function bases and explore the contributions of the additional time-delay terms.

It is worth noting that there are two parameters in the method: $k$ and $s$. The linear part of the function bases of the method consists of the input vector at the current and at $k-1$ previous times steps spaced by $s$, where $s-1$ is the number of skipped steps between consecutive observations. For the noiseless Lorenz system, $k = 2$ and $s = 1$ are often used. Increasing $k$ means adding more time-delay bases to the linear terms. And the $s$ is similar to the $\tau$ in the HOCC method, increasing $s$ is to train with data that skip more observations between consecutive observations. This section will optimize both parameters.

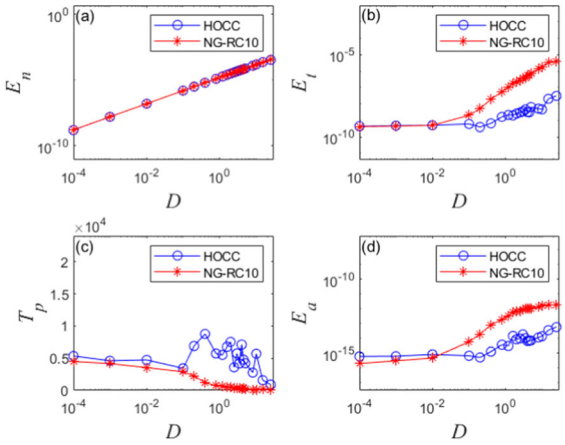


**Fig. 5** $E_n$, $E_t$, $T_p$, $E_a$ of the colored noise-driven system by NG-RC10 and HOCC methods. The blue line represents the HOCC method, and the red line represents the NG-RC10. (Color figure online)

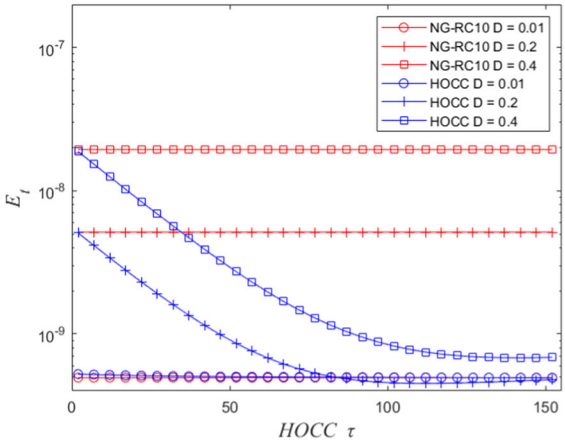


**Fig. 6** Relation between $E_t$ and $\tau$ of NG-RC10 and HOCC method. The sampling interval is $h = 0.0002$, and the noise correlation length is $\tau_d = 25h$

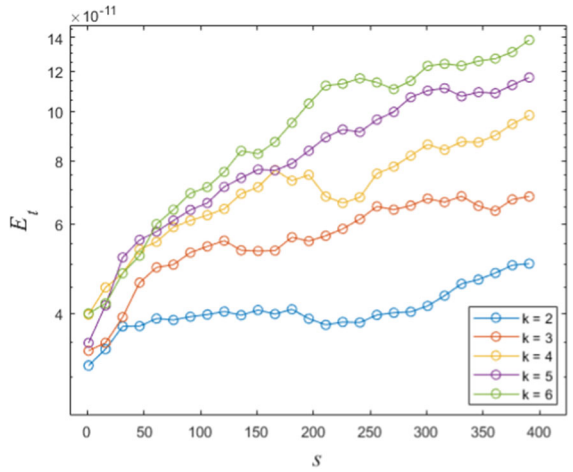


**Fig. 7** Relationship between $E_t$ and $k$ or $s$ of the white noise-driven system by NG-RC method with time-delay function bases. The parameters of the white noise-driven Lorenz system are $h = 0.0002$, $D = 0.1$

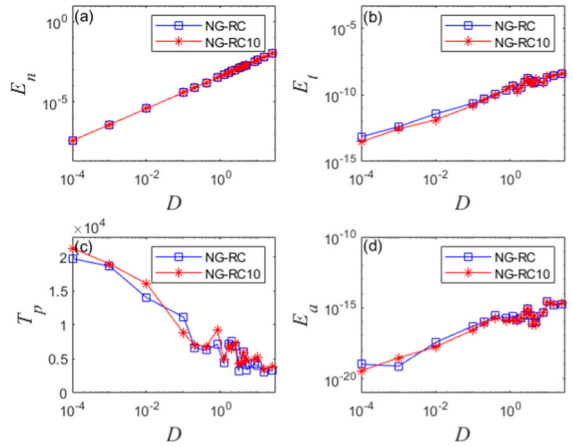


**Fig. 8** $E_t$,$E_n$,$T_p$,$E_a$ of the white noise-driven system by NG-RC method with time-delay function bases and NG-RC10. The blue line represents the NG-RC method with time-delay function bases, and the red line represents NG-RC10. (Color figure online)

## 4.1 White noise-driven system

To optimize parameters $k$ and $s$ to obtain the optimal prediction results, enough data was used for training. The $E_t$ under different $k$ and $s$ are obtained by training 20 white noise-driven systems with different $D$. Figure 7 shows the result when $D = 0.1$.

As shown in Fig. 7, $E_t$ increase with the increase of $k$ and $s$, indicating that increasing the number of time-delay bases in the basis function and using data after skipping more observations between consecutive observations for training are counterproductive to the noise resistance ability of the NG-RC method. Experiments on all the white noise-driven systems with different $D$ show similar results. The $k$ and $s$ that make $E_t$ reach the minimum are mostly $k = 2$ and $s = 1$, which are often used to predict the noiseless Lorenz system. With optimized parameters of $k$ and $s$, the results of the four indexes are shown in Fig. 8. See Supplementary Note 3 for the optimization process of $\alpha$.

Figure 8 shows the prediction results of NG-RC10 and NG-RC after optimized parameter in the white noise-driven system. The difference between NG-RC10 and NG-RC is small in all four numerical indicators. The NG-RC method with time-delay function bases has no significant effect on the prediction performance of the white noise-driven system. Hence both the NG-RC10 and NG-RC methods have a good white noise resistance.

### 4.2 Colored noise-driven system

As for the colored noise, the $k$ and $s$ are still optimized. Figure 9 shows the change of $E_t$ with $k$ and $s$ when $D = 0.0001$. For different $k$, when $s = 1$, $E_t$ is small, and then with the increase of $s$, $E_t$ rapidly peaks, and then gradually decreases. The difference is that the $E_t$ when $k = 2$ is always smaller than the $E_t$ corresponding to other $k$, and presents a U-shaped change with $s$, and its lowest point is close to the position where $E_t$ of HOCC method starts to stabilize when it changes with $\tau$ in Fig. 6. It can be seen that the selection of appropriate $k$ and $s$ in the time-delay bases have a certain effect on the denoising of the NG-RC method. Figure 10 shows the results of four indexes of the NG-RC method optimized after $k$ and $s$ and the NG-RC method without time-delay bases. Each data point in the figure is the result of $\alpha$ optimization, see Supplementary Note 4 for details.

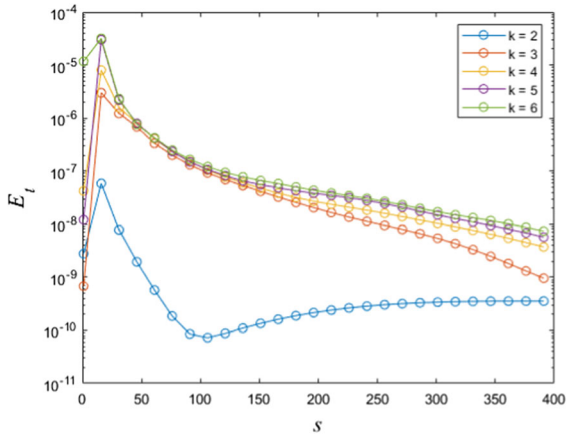


**Fig. 9** Relationship between $E_t$ and $k$ or $s$ of the colored noise-driven system by NG-RC method with time-delay function bases. The parameters of the colored noise-driven Lorenz system are $h = 0.0002$, $D = 0.0001$, and $\tau_d = 25h$

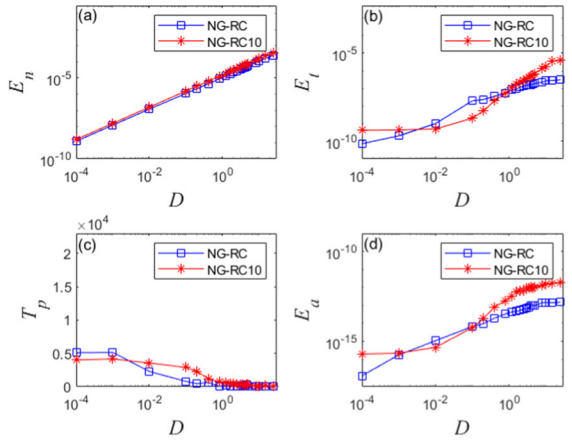


**Fig. 10** $E_n$, $E_t$, $T_p$, $E_a$ of the colored noise-driven system by NG-RC method with time-delay function bases and NG-RC10. The blue line represents the NG-RC method with time-delay function bases and the red line represents NG-RC10. (Color figure online)

In Fig. 10a, $E_n$ of the NG-RC method with time-delay bases is very similar to that of the NG-RC method without time-delay bases, indicating that they have similar fitting effects on training data. In Fig. 10b, the NG-RC method with time-delay bases is smaller than that without time-delay bases when the $D$ is relatively large or small. Similarly, for $T_p$ and $E_a$ in Fig. 10c, d, except for several data points with moderate $D$, the $T_p$ and $E_a$ of other data points of the NG-RC method with time-delay bases are smaller than those of NG-RC method without time-delay bases. It can be seen from the above results that the resistance ability to colored noise of the NG-RC method optimized with $k$, $s$, and $\alpha$ is improved compared with the NG-RC method without time-delay bases. However, compared with the HOCC method in Fig. 5, there is still a certain gap. Further improving the NG-RC's resistance ability to colored noise is an open question and a promising approach.

## 5 Conclusions

In this paper, the noise resistance ability of the NG-RC method is studied from the theoretical as well as numerical experimental, with reference to the HOCC method. Various aspects of these two methods are compared, such as the reconstruction error of parameters, training error $E_n$, prediction error $E_t$, prediction

time $T_p$, and the auto-correlation prediction error $E_a$. The similarity of the NG-RC and HOCC is shown theoretically, from which we study the NG-RC method without time-delay function bases. Such NG-RC10 method share the same function bases with HOCC. With different procedures and algorithms, we find that the NG-RC10 method has a better noise resistance when dealing with white noise than HOCC. Even for colored noise, the NG-RC10 method also shows a good prediction power when the noise strength is small, comparable with HOCC, while it cannot provides the correct reconstructed parameters.

The NG-RC method with time-delay function bases is also discussed. Such additional time-delay terms are helpful in the noise resistance of this method. The NG-RC method with and without time-delay function bases has the same noise resistance capacity for white noise. But for the colored noise, the NG-RC method with time-delay function bases works better by optimizing $k$ and $s$ in some conditions.

Reservoir computing methods, including the NG-RC method we discussed in this paper, is a promising study field of inverse problems of the dynamical system. In this paper, we show the simple ridge regression method in reservoir computing has a relatively strong noise resistance for white noise. But how to improve its noise resistance for colored noise is still an open question and a promising research topic.

**Author Contributions** All authors contributed to the study conception and design. Material preparation, numerical simulation and analysis were performed by SL and JG. The first draft of the manuscript was written by SL and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** This work is supported by the National Natural Science Foundation of China (NSFC) (Grant No. 11775034) and the Fundamental Research Funds for the Central Universities (Contract No. 2019XD-A10 and 2022RC26).

**Data Availability** Enquiries about data availability should be directed to the authors.

**Declarations**

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.



## References

1. Schütte, R., Zelewski, S.: Nonlinear Modeling And Forecasting. Addison-Wesley Publishing Company (1992)
2. Winkel, P.: Application of time series analysis in the clinical setting. Scand. J. Clin. Lab. Investig. **55**(s222), 11–16 (1995)
3. Denisse, P., Zbigniew, C., Benjamín, T.: Time series analysis in earthquake complex networks. Chaos **28**(8), 083128 (2018)
4. Gouveia, N.: Time series analysis of air pollution and mortality: effects by cause, age and socioeconomic status. J. Epidemiol. Community Health **54**(10), 750–755 (2000)
5. Caldarelli, G., Chessa, A., Pammolli, F., et al.: Reconstructing a credit network. Nat. Phys. **9**(3), 125–126 (2013)
6. Chen, G.R.: The China power grid: a network science perspective. Natl. Sci. Rev. **1**(3), 368 (2014)
7. Smith, A.: Genome sequence of the nematode c-elegans: a platform for investigating biology. Science **282**(5396), 2012–2018 (1998)
8. Beregi, S., Barton, D., Rezgui, D., et al.: Robustness of nonlinear parameter identification in the presence of process noise using control-based continuation. Nonlinear Dyn. **104**, 885–900 (2021)
9. Yamakou, M.E., Jost, J.: Coherent neural oscillations induced by weak synaptic noise. Nonlinear Dyn. **93**, 2121–2144 (2018)
10. Hametner, C., Kozek, M., Böhler, L., et al.: Estimation of exogenous drivers to predict COVID-19 pandemic using a method from nonlinear control theory. Nonlinear Dyn. **106**, 1111–1125 (2021)
11. Costa, J., Liu, K., So, H.C., et al.: Multidimensional prewhitening for enhanced signal reconstruction and parameter estimation in colored noise with Kronecker correlation structure. Signal Process. **93**(11), 3209–3226 (2013)
12. Huang, J., Hu, W.D., Du, X.Y., et al.: Parameter estimation for space surveillance based on sparse reconstruction. In: 2012 3rd International Workshop on Cognitive Information Processing (CIP), pp. 1–6. (2013)
13. Koda, M., Seinfeld, J.H.: Reconstruction of atmospheric pollutant concentrations from remote sensing data-an application of distributed parameter observer theory. IEEE Trans. Autom. Control **27**(1), 74–80 (2003)
14. Li, H.P.: Analysis of reconstruction method of dynamics system under measurement noise. Master Thesis, Bejing University of Posts and Telecommunications (2019)

15. Lu, J.N., Lu, J.H., Xie, J., et al.: Reconstruction of the Lorenz and Chen systems with noisy observations. Comput. Math. Appl. **48**(8–9), 1427–1434 (2003)
16. Wang, J., Yan, Z., Gui, L., et al.: Reconstruction of nonlinear flows from noisy time series. Nonlinear Dyn. **108**(4), 3887–3902 (2022)
17. Zhang, Z.Y., Zheng, Z.G., Niu, H.J., et al.: Solving the inverse problem of noise-driven dynamic networks. Phys. Rev. E **91**, 012814 (2015)
18. Levnaji, Z., Pikovsky, A.: Untangling complex dynamical systems via derivative-variable correlations. Sci. Rep. **4**, 18 (2014)
19. Han, X., Shen, Z., Wang, W.X., et al.: Robust reconstruction of complex networks from sparse data. Phys. Rev. Lett. **114**, 028701 (2015)
20. Wu, X., Wang, W., Wei, X.Z.: Inferring topologies of complex networks with hidden variables. Phys. Rev. E **86**, 046106 (2012)
21. Wang, W.X., Yang, R., Lai, Y.C., et al.: Predicting catastrophes in nonlinear dynamical systems by compressive sensing. Phys. Rev. Lett. **106**, 154101 (2011)
22. Brunton, S.L., Proctor, J.L., Kutz, J.N.: Discovering governing equations from data: sparse identification of nonlinear dynamical systems. Proc. Natl. Acad. Sci. USA **113**, 3932 (2015)
23. Lukoševičius, M.: A practical guide to applying echo state networks. Lecture Notes in Computer Science (2012)
24. Jaeger, H.: Harnessing nonlinearity: predicting Chaotic systems and saving energy in wireless communication. Science **304**(5667), 78–80 (2004)
25. Manjunath, G., Jaeger, H.: Echo state property linked to an input: exploring a fundamental characteristic of recurrent neural networks neural computation. Neural Comput. **25**(3), 671–696 (2013)
26. Chen, T.Y., Chen, Y., Yang, H.J., et al.: Reconstruction of dynamic structures of experimental setups based on measurable experimental data only. Chinese Physics B. **27**(3), 030503 (2018)
27. Lorenz, E.N.: Deterministic nonperiodic flow. J. Atmos. Sci. **20**(2), 130–141 (1963)
28. Jaeger, H.: The "echo state" approach to analysing and training recurrent neural networks-with an erratum note'. German National Research Center for Information Technology GMD Technical Report, Bonn. **148**(34), 13 (2001)
29. Maass, W., Natschlager, T., Markram, H.: Real-time computing without stable states: a new framework for neural computation based on perturbations. Neural Comput. **14**(11), 2531–2560 (2002)
30. Peng, Y., Wang, J.M., Peng, X.Y.: Survey on reservoir computing. Acta Electron. Sin. **39**(010), 2387–2396 (2011)
31. Verstraeten, D., Schrauwen, B., Haene, M.D., et al.: An experimental unification of reservoir computing methods. Neural Netw. **20**(3), 391–403 (2007)
32. Kong, L.W., Fan, H.W., Grebogi, C., et al.: Machine learning prediction of critical transition and system collapse. Phys. Rev. Res. **3**(1), 013090 (2021)
33. Kong, L.W., Fan, H.W., Grebogi, C., et al.: Emergence of transient chaos and intermittency in machine learning. J. Phys. Compl. **2**(3), 035014 (2021)
34. Jiang, J., Huang, Z.G., Grebogi, C., et al.: Predicting extreme events from data using deep machine learning: when and where. Phys. Rev. Res. **4**, 023028 (2022)
35. Gauthier, D.J., Bollt, E., Griffith, A., et al.: Next generation reservoir computing. Nat. Commun. **12**, 5564 (2021)
36. Chen, Y., Zhang, C.Y., Chen, T.Y., et al.: Reconstruction of noise-driven nonlinear dynamic networks with some hidden nodes. Sci. China Phys. Mech. Astron. **60**(7), 8 (2017)
37. Chen, Y.: Structures and functions self-organized explosive synchronization in complex network and network reconstructions with strong noises. Ph.D. Thesis, Bejing University of Posts and Telecommunications (2018)
38. Guo, Y.L., Zhang, H., Wang, L., et al.: Transfer learning of chaotic systems. Chaos **31**, 011104 (2021)