2020 International Symposium on Nonlinear Theory and Its Applications,
NOLTA2020, Virtual, November 16-19, 2020

Echo State Network with Chaos Noise for Time Series Prediction

Yoko Uwate†, Martin Schule ‡, Thomas Ott ‡ and Yoshifumi Nishio†

†Dept. of Electrical and Electronic Engineering, Tokushima University
2-1 Minami Josanjima, Tokushima, Japan
Tel: +81-88-656-7662, Fax: +81-88-656-7471
Email: {uwate, nishio}@ee.tokushima-u.ac.jp
‡Institute of Applied Simulation, Zurich University of Applid Sciences
Einsiedlerstrasse 31a, 8820 Waedenswil, Switzerland
Email: {scli, thomas.ott}@zhaw.ch

Abstract—In this study, performance of chaos noise injected to Echo State Network for time series prediction is investigated. For the evaluation of the chaos noise, two parameters of the logistic map are selected to produce different features as intermittency chaos and fully developed chaos. By computer simulations, it is confirmed that the three-periodic intermittency chaos noise is better performance than the fully developed chaos noise for time series prediction.

1. Introduction

Today, every time we open a newspaper, the page contains AI characters, and new technologies that apply deep learning are being developed everywhere in the industry. However, deep neural networks have problems such as requirement of enormous amount of training data and difficult to implement in hardware. In recent years, research on reservoir computing has been actively investigated in order to facilitate hardware implementation and reduce the calculation cost. The network structure of reservoir computing is composed of an input layer, a hidden layer called a reservoir, and an output layer. Representative examples of reservoir layers that play a particularly important role in learning are the Echo State Network (ESN) [1], [2], in which the neuron network is random and the weight is fixed, and the Liquid State Computing (LSC) [3], which uses the spiking neuron model.

Jaeger *et al.* have proposed adding the state noise to the update equation of ESNs [4]. They confirm that adding state noise is computationally more expensive, but appears to have the additional benefit of stabilizing solutions in models with output feedback. Hayakawa *et al.* pointed out the chaos near the three-periodic window of the logistic map gains best performance of Hopfield Neural Networks (HNNs) for solving traveling salesman problems (TSPs) [5]. In our previous studies, we have proposed a method of modeling chaos noise to prove the effectiveness of chaos noise, and confirmed that the noise obtained by modeling shows the same performance as chaos noise [6], [7]. Therefore, we consider that chaotic noise is effective for ESNs.

We investigate the performance of chaos noise when it is injected to the ESN for time series prediction. Computer simulated results show that the intermittency chaos noise gains good performance for long-term prediction of time series.

2. Echo State Network

Figure 1 shows an echo state network. ESNs are applied to supervised machine learning tasks where for a given training input signal $\mathbf{u}(n)$ and a desired target output signal $\mathbf{y}^{target}(n)$ is known. The task of ESN is to minimize the difference between the output value $\mathbf{y}(n)$ and the target value $\mathbf{y}^{target}(n)$. The error E is calculated by the following equation which called a Mean-Square Error (MSE).

$$E(\mathbf{y}, \mathbf{y}^{target}) = \frac{1}{T}\sum_{n=1}^{T}(y_i(n) - y_i^{target}(n))^2 \tag{1}$$

where T is the number of data points in the training dataset.

The update equation of ESNs is described as follows.

$$\begin{aligned} \tilde{\mathbf{x}}(n) &= \tanh(\mathbf{W}^{in}[1; u(n)] + \mathbf{W}\mathbf{x}(n-1)), \\ \mathbf{x}(n) &= (1-\varepsilon)\mathbf{x}(n-1) + \varepsilon\tilde{\mathbf{x}}(n) \end{aligned} \tag{2}$$

where $\mathbf{x}(n)$ is a vector of reservoir neuron activations and $\tilde{\mathbf{x}}(n)$ is its update with all at time step n and ε denotes the leaking rate.

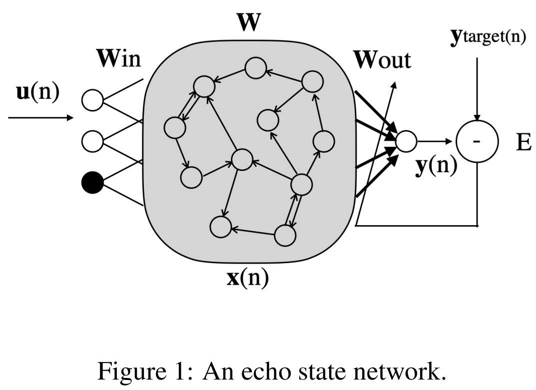


Figure 1: An echo state network.

For the computer simulations, the parameter of the number of neurons in reservoir is set to 1,000 and the leaking rate (ε) is fixed with 0.3.

3. Chaos Noise

In this section, we explain chaos noise injected to the ESN. The logistic map is used to generate the chaos noise:

$$v(n+1) = \alpha v(n)(1 - v(n)). \tag{3}$$

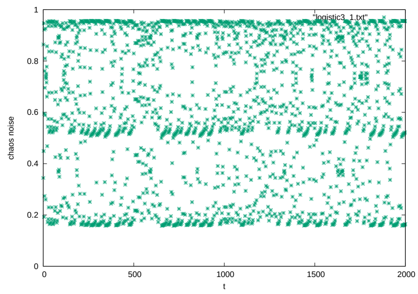
Varying the parameter α, Eq. (3) behaves chaotically via a period-doubling cascade. Further, it is well known that the map produces intermittent chaos just before periodic windows appear.

Figure 2 (a) shows an example of the intermittency chaos near the three-periodic window obtained from Eq. (3) for $\alpha = 3.8274$. As we can see from the figure, the chaotic time series could be divided into two phases; laminar parts of periodic behavior with period three and burst parts of spread points over the invariant interval. As increasing α, the ratio of the laminar parts becomes larger and finally the three-periodic window appears. For the comparison, we also carry out computer simulations for the case of fully developed chaos in Fig. 2 (b) for $\alpha = 4.0000$. This noise is much more similar to the random noise.

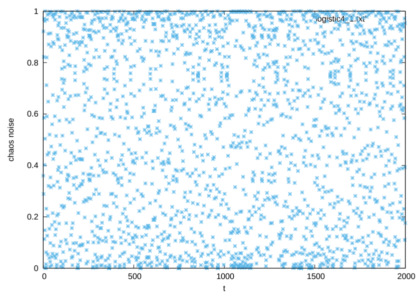
The state equation of ESNs with chaos noise is described by the following equation.

$$\mathbf{x}(n) \quad = \quad (1-\varepsilon)\mathbf{x}(n-1) + \varepsilon\tilde{\mathbf{x}}(n) + \delta\mathbf{v}(n). \tag{4}$$

where δ is the control parameter of the amplitude of the chaos noise. For the simulation, δ is set to 0.0001.



(a) 3-periodic intermittency chaos (α=3.8274).



(b) Fully developed chaos (α=4.0000).

Figure 2: Two types of chaos noise generated from logistic map.

4. Time Series Prediction

We use the Mackey-Glass equation for the time series prediction. The Mackey-Glass equation is the representative example of delay induced chaotic behavior. The equation is the nonlinear time delay differential equation.

$$\frac{dz}{dt} = \beta \frac{z_\tau}{1 + z_\tau^m} - \gamma z, \qquad \gamma, \beta, m > 0 \tag{5}$$

An example of time series of Mackey-Glass is shown in Fig. 3.

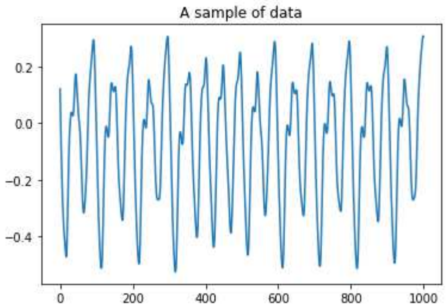


Figure 3: An example of time series of Mackey-Glass equation ($\tau = 17$).

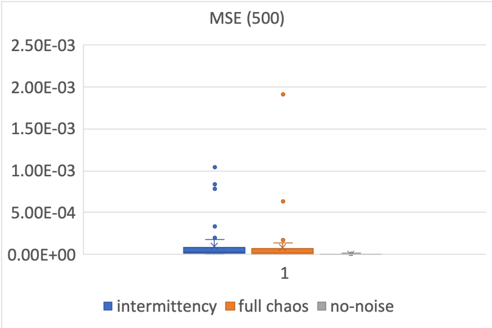
For the training of ESNs, the number of points of time series is set to 2,000. After the training, 2,000 time series points will be predicted with free running of ESNs. For implementation of ESNs, we referred the manuscript [8].

5. Simulation Results

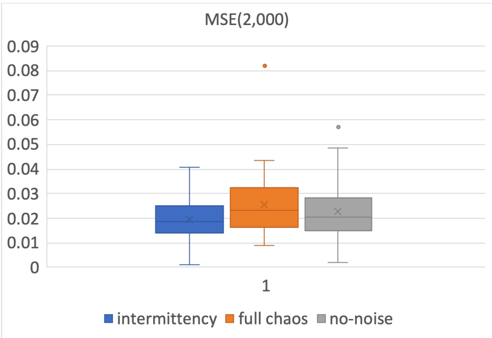
The computer simulation results of 50 time average are shown in Fig. 4. We calculate two types of MSE after training. The first one is that MSE is calculated for the first 500 time steps. The second one is that MSE is calculated for the whole 2,000 time steps. Figure 4 (a) shows the simulation results of MSE with first 500 time steps. From this figure, we can see that the ESN without noise has the lowest MSE. The performance of intermittency chaotic noise and fully developed chaos noise of MSEs are almost the same.

Figure 4 (b) shows the simulation results of MSE with 2,000 time steps. From this figure, it can be seen that in the case of ESN without noise, the range of values taken by MSE is the largest and fully developed chaos is the smallest. ESNs using intermittency chaos noise has a lower range of values that MSE can take than other ESNs.

Next, we summarize the 50 times average results of MSE using three types of ESNs, ESN with intermittency chaos noise, ESN with fully developed chaos noise and ESN without noise (Table 1). This result is very interesting. This is because ESN with no noise shows the best performance in MSE for first 500 steps, whereas ESN with intermittency chaotic noise shows the best accuracy in MSE for whole 2,000 steps. In other words, the accuracy immediately after learning is good for ESN without noise, however it can be said that ESN with intermittency chaos noise is effective for long prediction.



(a) MSE (500).



(b) MSE (2,000).

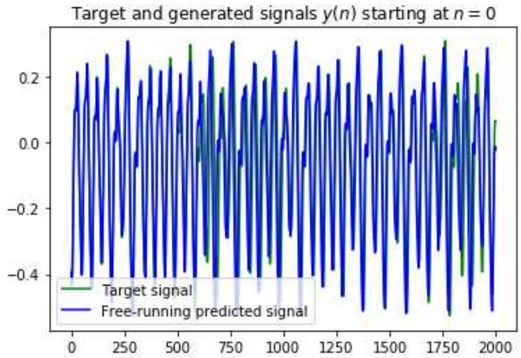
Figure 4: Simulation results of MSE.

Table 1: Simulation results (50 times average)

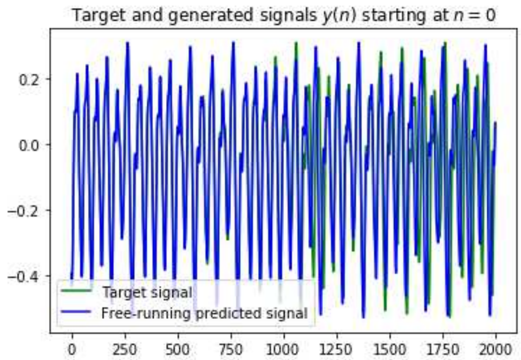
	MSE (500)	MSE (2,000)
Intermittency chaos	1.02e-04	1.96e-02
Fully chaos	9.04e-05	2.55e-02
No noise	6.43e-06	2.25e-02

Figure 5 shows the example of simulation results of target data and predicted data after the training of ESNs. In the case of ESN without noise, the target and predicted value are in good agreement up to the first 500 steps, but after that, the error becomes large. While, in the case of ESN with intermittency chaos noise, it can be seen that the error is small over 2000 steps.

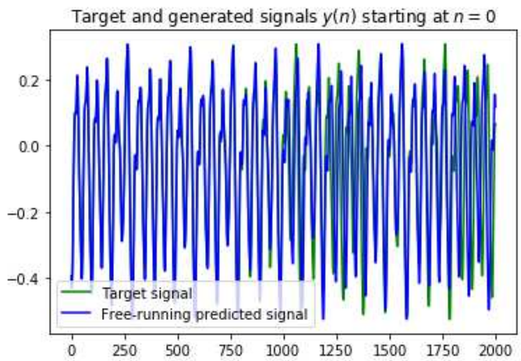
Finally, the obtained output weights ($\mathbf{W}^{out}$) of ESNs are shown in Fig. 6 It seems that the ESN with chaos noise has a higher ($\mathbf{W}^{out}$) value than the ESN without noise. A more detailed study of the learning process of chaos noise and the performance of ESNs is one of our future topics.



(a) ESN with 3-periodic intermittency chaos (α=3.8274).

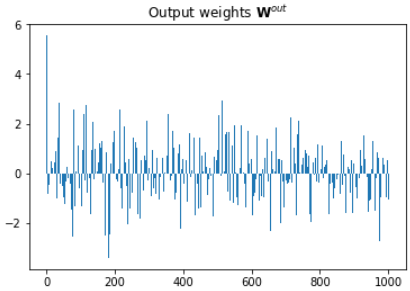


(b) ESN with Fully developed chaos (α=4.0000).

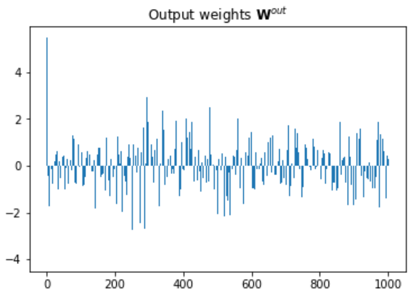


(c) ESN without state noise.

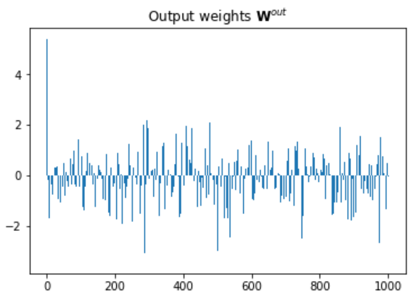
Figure 5: Simulation results of time series prediction using three types of ESNs.



(a) ESN with 3-periodic intermittency chaos (α=3.8274).



(b) ESN with Fully developed chaos (α=4.0000).



(c) ESN without state noise.

Figure 6: Output weights after training.

6. Conclusions

When using ESN in real-world nonlinear modeling tasks, the ultimate goal is to minimize test errors. The adding state nose to ESN has been proposed to have the benefit of stabilizing solutions in models with output feedback. In this study, we investigated the performance of chaos noise when it is injected to the ESN for time series prediction. Computer simulated results showed that the intermittency chaos noise gains good performance for long-term prediction of time series.

In future works, we would like to make clear the reason why intermittency chaos noise better performance than the others for time series prediction. Furthermore, we would like to extend this study and investigate reservoir computing using chaotic circuits in the reservoir layer [9], [10].

References

[1] H. Jaeger, ""Echo State" Approach to Analyzing and Training Recurrent Neural Networks," *Technical Report GMD Report 148*, German National Research Center for Information Technology, 2001.

[2] H. Jaeger, "Echo State Network," *Scholarpedia*, 2(9):2330, 2007.

[3] M. Lukosevicius, "A Practical Guide to Applying Echo State Networks," *Neural Networks: Tricks of the Trade*, 7700, pp. 659-686, 2012.

[4] H. Jaeger, M. Lukosevicius, D. Popovici and U. Siewert, "Optimization and Applications of Echo State Networks with Leaky-integrator Neurons," *Neural Networks*, 20, pp. 335352-686, 2007.

[5] Y. Hayakawa and Y. Sawada, "Effects of the Chaotic Noise on the Performance of a Neural Network Model for Optimization Problem," *Phys. Rev. E, State. Phys. Plasmas Fluids Relat. Interdiscip Top.*, vol. 51, no. 4, pp.2693-2696, 1995.

[6] Y. Uwate, Y. Nishio and A. Ushida, "Markov Chain Modeling of Intermittency Chaos and its Application to Hopfield NN," *IEICE Transactions on Fundamentals*, vol. E87-A, no. 4, pp. 774-779, Apr. 2004.

[7] Y. Uwate, Y. Nishio, T. Ueta, T. Kawabe and T. Ikeguchi, "Performance of Chaos and Burst Noises Injected to the Hopfield NN for Quadratic Assignment Problems," *IEICE Transactions on Fundamentals*, vol. E87-A, no. 4, pp. 937-943, Apr. 2004.

[8] M. Lukosevicius, "A Practical Guide to Applying Echo State Networks," *Neural Networks: Tricks of the Trade*, 7700, pp. 659-686, 2012.

[9] T. Yamane, Y. Katayama, R. Nakane, G. Tanaka and D. Nakano, "Wave-based Reservoir Computing by Synchronization of Coupled Oscillators," *Proceedings of International Conference on Neural Information Processing (ICONIP'15)*, pp. 198-205, 2015.

[10] J. Choi and P. Kim, "Reservoir Computing based on Quenched Chaos," *Chas, Solitons and Fractals*, 140, 110131, https://doi.org/10.1016/j.chaos.2020.110131, 2020.