Forecasting the chaotic dynamics of external cavity semiconductor lasers

Kai, Chao; Li, Pu; Yang, Yi; Wang, Bingjie; Shore, K. Alan; Wang, Yuncai

Optics Letters

DOI:
10.1364/OL.480874

Published: 25/02/2023

Forecasting the chaotic dynamics of external cavity semiconductor lasers

CHAO KAI,[1] PU LI,[1,2,*] YI YANG,[1] BINGJIE WANG,[1] K. ALAN SHORE,[3] AND YUNCAI WANG[2]

[1]Key Laboratory of Advanced Transducers and Intelligent Control System, Ministry of Education, College of Physics and Optoelectronics, Taiyuan University of Technology, Taiyuan 030024, China
[2]Guangdong Provincial Key Laboratory of Photonics Information Technology, School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China
[3]School of Electronic Engineering, Bangor University, Wales LL57 1UT, UK
**Corresponding author: lipu8603@126.com*

Received XX Month XXXX; revised XX Month, XXXX; accepted XX Month XXXX; posted XX Month XXXX (Doc. ID XXXXX); published XX Month XXXX

Chaotic time series prediction has been paid intense attention in recent years due to its important applications. Herein, we present a single-node photonic reservoir computing approach to forecast the chaotic behavior of external cavity semiconductor lasers using only observed data. In the reservoir, we employ a semiconductor laser with delay as the sole nonlinear physical node. Through investigating the effect of the reservoir meta-parameters on the prediction performance, we numerically demonstrate that there exists an optimal meta-parameter space for forecasting optical-feedback-induced chaos. Simulation results demonstrate that using our method, the upcoming chaotic time series can be continuously predicted for a time period in excess of 2 ns with a normalized mean squared error lower than 0.1. This proposed method only utilizes simple nonlinear semiconductor lasers and thus offers a hardware-friendly approach for complex chaos prediction. In addition, this work may provide a roadmap for the meta-parameter selection of a delay-based photonic reservoir to obtain optimal prediction performance.

Forecasting the dynamics of chaotic systems has important applications such as security analysis for random number generators (RNG) [1,2] and chaos synchronization in private communications [3,4]. The most common forecasting approach is based on regression models [5]. Unfortunately, these model-driven methods require strong empirical regularities of chaotic systems that are usually difficult or even impossible to be obtained in practice [6].

In recent years, artificial neural networks (ANNs) based methods have attracted increasing attention to perform chaos prediction, because they only use observed data to execute the prediction [6,7]. Generally, ANNs based methods can be divided into two types depending on their different topologies. One is based on feedforward networks (FFNs), which are mostly applied to static tasks such as recognition and detection but not for prediction due to their feedforward structure [8]. The other is based on recurrent neural networks (RNNs). This kind of ANNs is more suitable for forecasting tasks because their neurons have self-loops and backward connections. For example, Logar *et al.* utilized a RNN employing three fully connected layers and forty self-loops to estimate the chaotic output of Mackey-Glass equation [9]. Zhang *et al.* realized multi-step prediction of Lorenz and Mackey-Glass equations by employing RNNs changing 5000 weights in each training epoch [10]. Nevertheless, these traditional RNNs have to train each weight in a fully connected topology with at least three layers, so that their training processes have the problems of exploding or vanishing gradients and demand extensive computational power [7, 8]. That is also the main reason why most of these traditional RNNs can only be applied to predict low-dimensional chaos from simple mathematical models or electronic circuits [8, 11, 12].

In fact, there is a kind of ANN - reservoir computing (RC)-which can remove the aforementioned challenges, because RCs only need to train the output layer and thus requires minimal computational resources [7,8]. Until now, there have been few of reports using RC to predict high-dimensional chaotic behaviors such as spatio-temporal chaos and optical chaos [7, 11-13]. Typically, Pathak *et al.* utilized parallel RC with 10000 nodes to forecast the behavior of a large spatio-temporally chaotic system [12]. Ailm *et al.* predicted the peak amplitudes of chaotic laser pulses from semiconductor lasers with optical injection using RC with 6000 hidden nodes [13]. Li *et al.* employed RC with 5000 nodes to forecast the intensity time series of an optically injected semiconductor laser over a time duration of 0.6 ns [11]. Although using RC has led to some remarkable achievements, one must notice that the conventional RC used in the above reports still have thousands of physical nodes and need the help of a high-powered computer. Thus, they suffer from having inefficient implementations on electronic hardware [14].

In this letter, we propose and demonstrate a simple method utilizing a single-node-based RC to implement high-dimensional optical chaos prediction. Such a simple reservoir structure constructed with a single nonlinear physical node enable our method to reduce the complexity of the whole forecasting system and thus incurs little computational cost. Specifically, we employ a photonic RC (PRC) based on a single node to forecast the chaotic dynamics of external cavity semiconductor lasers (ECSL). In this method, the photonic reservoir is constructed with a single nonlinear physical node (*viz.*, an optically injected semiconductor laser with a time delay loop). Through optimizing the effect of the reservoir meta-parameters such as injection strength, frequency detuning and feedback strength on the prediction performance, we successfully forecast a continuous chaotic time series over 2 ns by using only the observed data. Comparing with previous reports, the merits of our work lie in: (i) This work extends the applications of the single-node PRC, whose early focus was on speech recognition [15-17], nonlinear channel equalization [15,18,19] and radar signal forecasting [18,19]. It is the first time to our knowledge that a single-node PRC has been utilized to forecast a high-dimensional chaos laser. (ii) In contrast to previous reports on laser chaos forecasting, the chaotic behaviors predicted here are generated by an ECSL, rather than an optically injected laser. Usually, optically injected lasers only have six variables and thus a relatively low dimensionality of chaos [20]. In comparison, the ECSL is infinite dimensional in its state space [20,21], and thus is difficult to be tackled from the mathematical side. Moreover, it is also of significance to forecast the chaotic dynamics of the ECSL from a practical perspective, because it is the most widely used optical chaos source [22-25].

Figure 1 sketches the operating principle of our single-node PRC for chaos prediction. The single-node PRC includes three layers: the input layer, the reservoir layer, and the output layer. In the input layer, chaotic time series $I_{in}(t)$ generated from an ECSL is multiplied by a mask vector $\boldsymbol{M(t)}$ and an input gain G_{in} to obtain the input vector $\boldsymbol{S(t)}$:

$$\boldsymbol{S(t)}=I_{in}(t)\times \boldsymbol{M(t)}\times G_{\text{in}}, \tag{1}$$

Note, $\boldsymbol{M(t)}$ is a six-level random vector {±0.5, ±1, ±1.5} with a length of T, which is equivalent to the input connection weights of traditional reservoirs. G_{in} is a factor that achieves the linear scaling of the input vector $\boldsymbol{S(t)}$.

In the reservoir layer, the drive laser (D-L) is modulated by the input vector $\boldsymbol{S(t)}$ via a phase modulator (Mod) and then injected into the response laser (R-L) with a time delay loop. This delay loop contains N equidistant points separated in time by $\theta = \tau/N$, where τ is the loop length. These N equidistant points are so-called "virtual nodes" since their roles are similar to that of the nodes in traditional reservoirs. Therefore, the state $X_i(t)$ of each virtual node corresponds to a transient response of R-L within an interval θ. Within the whole feedback delay time τ, these N states of virtual nodes make up a vector $\boldsymbol{X_i(t)}$, (i = 1, 2, ..., N). In this way, the input data $\boldsymbol{S(t)}$ is nonlinearly mapped into N-dimensional space. The operation process in the reservoir can be modelled as follows:

$$\begin{aligned}\frac{dE(t)}{dt} &= \frac{1+i\alpha}{2}\left\{\frac{g[N(t)-N_0]}{1+\varepsilon|E(t)|^2}-\frac{1}{\tau_{\text{p}}}\right\}E(t)\\ &+\frac{k_{\text{f}}}{\tau_{\text{in}}}E(t-\tau)\exp(-i2\pi\nu\tau)\\ &+\frac{k_{\text{inj}}}{\tau_{\text{in}}}E_{\text{inj}}(t)\exp(i2\pi\Delta\nu t)+\sqrt{2\beta N(t)}\chi(t),\end{aligned} \tag{2}$$

$$\frac{dN(t)}{dt} = J-\frac{N(t)}{\tau_{\text{s}}}-\frac{g[N(t)-N_0]}{1+\varepsilon|E(t)|^2}|E(t)|^2, \tag{3}$$

$$E_{\text{inj}}(t)=\sqrt{I_{\text{d}}}e^{(i\pi s(t))}, \tag{4}$$

where $E(t)$ and $N(t)$ are the complex electric field amplitude and carrier density of the R-L, respectively. $E_{inj}(t)$ is the complex electric field amplitude of the modulated D-L. k_{inj} and k_f represent the injection strength from D-L to R-L and the feedback strength of the R-L, respectively. ν is the frequency of the free-running R-L, while $\Delta\nu$ denotes the frequency detuning between the D-L and the R-L. I_d is the optical intensity of the D-L. The other parameters and their values in our simulation are as follows: linewidth enhancement factor α = 5.0, carrier density at transparency N_0 = 1.4×10^{24} m^{-3}, differential gain coefficient g = 1.414×10^{-12} m^3·s^{-1}, gain saturation coefficient ε = 5.0×10^{-23}, internal

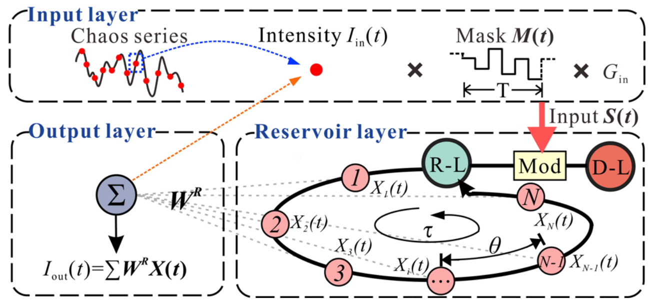


Fig. 1. Schematic of single-node PRC for chaos prediction. D-L is the drive laser. R-L is the response laser. Mod is the phase modulator.

cavity round trip time τ_{in} = 7.38 ps, photon lifetime τ_p = 1.92 ps, carrier lifetime τ_s = 2.04 ns, R-L injection current J = 18 mA and its feedback delay time τ = 10 ns. In addition, $\chi(t)$ is the Gaussian noise with zero mean and unity variance and β is the noise strength set as its typical value of 4.5×10^{-4} [26].

In the output layer, the output $I_{out}(t)$ for the input data I_{in}(t) is calculated as a linear combination of the state vector of virtual nodes $\boldsymbol{X_i(t)}$ with the output connection weights vector $\boldsymbol{W^R}$ for every temporal periodicity T, as shown in Eq. (5). We use the ridge regression algorithm to train the output connection weights vector $\boldsymbol{W^R}$, as shown in Eq. (6). In the training procedure, the ridge parameter λ is set to its typical value 10^{-6} [27]. Through the training phase, $I_{out}(t)$ approaches the target $I_{target}(t)$ as closely as possible.

$$I_{out}(t)=\sum_{i=1}^{N}\boldsymbol{W_i}^R\boldsymbol{X_i(t)} \tag{5}$$

$$\boldsymbol{W}^R = \arg\min\left\|I_{\text{target}}(t)-I_{\text{out}}(t)\right\|^2+\lambda\left\|\boldsymbol{W}^R\right\|^2 \tag{6}$$

After the training is completed, the trained vector $\boldsymbol{W^R}$ remain constant. Then, we continuously use the output $I_{out}(t)$ as the next input data $I'_{in}(t)$ in the input layer to forecast the future chaotic behaviors, as indicated by the orange dashed line in Fig. 1. The prediction performance is evaluated by calculating the normalized mean squared error (NMSE) between the prediction result and the target value [28]. Because we employ five sets of data for the prediction, their average of NMSEs in five sets is the final evaluation result. The NMSE of each data set is defined as below:

$$NMSE=\frac{1}{L}\frac{\sum_{t=1}^{L}(I_{\text{out}}(t)-I_{\text{target}}(t))^2}{\text{var}<I_{\text{out}}>}, \tag{7}$$

where $I_{out}(t)$ stands for the predicted results of the single-node-based PRC. $I_{target}(t)$ denotes the target value. L is the total number of test data and var<·> presents the variance.

In the following, we will validate the feasibility of our method. First, we produce chaotic intensity time series through numerical simulation for training and testing the PRC. Here, chaotic dynamics in the ECSL is described using the Lang-Kobayashi rate equations [29], which consists of Eqs. (2) and (3) modified as follows: (i) the injection term $k_{inj}/\tau_{in}E_{inj}(t)\exp(i2\pi\Delta\nu t)$ in Eq. (2) is omitted; (ii) some critical parameters are reset. That is, feedback strength k_f' = 0.054, the feedback delay time τ' = 1 μs, and the noise strength β = 1.5 × 10^{-6}. In this simulation, 10^5 data points of chaotic intensity time series are finally recorded with a sampling period of 10 ps. Figure 2(a) shows the power spectrum of the chaotic time series with a relaxation oscillation frequency about 4 GHz, which is similar with that in an experimental ECSL. Further, we can confirm from its associated phase diagram [Fig. 2(b)] that the trajectories of attractor are complicated and cover a wide area, which is more complex and unordered than the chaotic attractor of the optically injected laser as shown in Ref. [11].

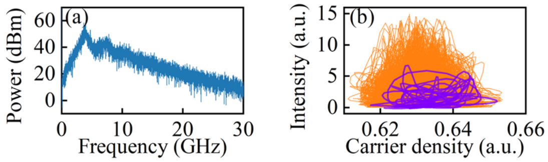


Fig. 2. (a) Power spectrum and (b) Phase diagram of the ECSL with a feedback delay time of 1 µs. The orange and the violet attractors are plotted for the 500 ns and the 5 ns chaotic time series, respectively.

Next, we determine the optimal number of training samples N_s, number of virtual nodes N_v, input gain G_{in} and leakage rate δ of the PRC one by one. To begin, we initialize these four parameters as follows: N_s = 2500, N_v = 1000, G_{in} = 1 and δ = 0.15. After that, we first optimize the number of samples N_s while keeping the other three parameters fixed. From the associated result [Fig. 3(a)], we confirm that N_s = 15000 is the best compromise between the prediction performance and the training efficiency, where the corresponding NMSE is 1.47. Secondly, we optimize the number of virtual nodes N_v. In PRC, the number of virtual nodes determine whether the reservoir can fully learn the characteristics of the training samples. Figure 3(b) depicts the dependence of the prediction ability on the number of virtual nodes N_v, where N_s is fixed at its optimal value 15000 while G_{in} and δ are fixed at their own initial values. From Fig. 3(b), we find that the NMSE decreases with increase of N_v and tends to a stationary value. Thus, we choose N_v = 800 where the NMSE corresponds to the lowest value 1.44. Thirdly, we determine the input gain G_{in}, which is used to adjust the amplitude of the input signal. As shown in Fig. 3 (c), when N_s = 15000, N_v = 800 and δ = 0.15, a relatively small NMSE = 0.86 at the case of G_{in} = 1.5. Further, we optimize the last parameter (i.e., the leakage rate δ) by setting the other three parameters at their optimal values (N_s = 15000, N_v = 800 and G_{in} = 1.5). δ contributes to the update speed of the reservoir by establishing a connection among every two virtual nodes. The smaller δ is, the faster the update speed. From Fig. 3(d), it can be clearly seen that the optimal leakage rate δ is 0.25 because the associated NMSE at this point reaches its lowest value 0.59. Considering the numerous training samples and relatively few virtual nodes in the training phase, an appropriately fast update speed is in line with practical needs.

Then, we consider the effect of critical meta-parameters of the reservoir, which are the injection strength k_{inj}, frequency detuning Δν and the feedback strength k_f, respectively. (i) Figures 4(a1) and (a2) show the effect of k_{inj} on the dynamics and performance of the PRC, respectively. k_{inj} mainly affects the consistency of transient response, which is a fundamental feature of RC and relates to the reproducibility of system responses under repetitive injection of similar inputs [14]. From the bifurcation diagram of the free running system [Fig. 4(a1)], we can find that when k_{inj} is small, the dynamics of system is a single-cycle

Fig. 3. NMSE versus the (a) number of training samples N_s for N_v = 1000, G_{in} = 1 and δ = 0.15. (b) number of virtual nodes N_v for N_s = 15000, G_{in} = 1 and δ = 0.15. (c) input gain G_{in} for N_s = 15000, N_v = 800 and δ = 0.15. (d) leakage rate δ for N_s = 15000, N_v = 800 and G_{in} = 1.5.

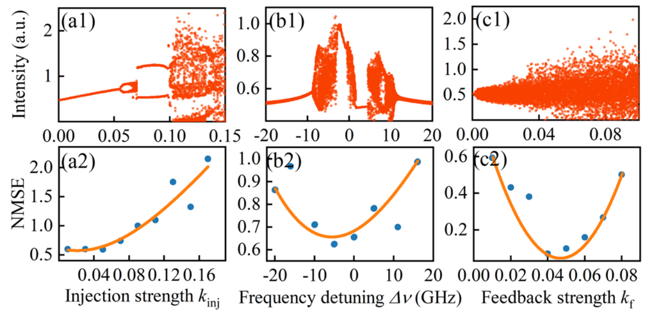
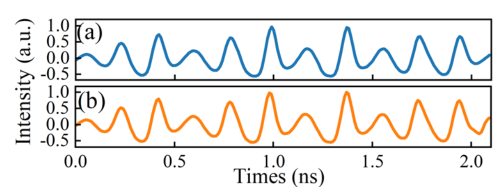


Fig. 4. (a1) Bifurcation diagram of the R-L output and (a2) NMSE versus the injection strength k_{inj} for k_f = 0.01 and $\Delta\nu$ = -5 GHz. (b1) Bifurcation diagram of the R-L output and (b2) NMSE versus the frequency detuning $\Delta\nu$ for k_f = 0.01 and k_{inj} = 0.05. (c1) Bifurcation diagram of the R-L output and (c2) NMSE versus the feedback strength k_f for k_{inj} = 0.05 and $\Delta\nu$ = -5 GHz.

state. With further increase of k_{inj}, the dynamical state of the PRC becomes more complex. Combined with Fig. 4 (a2), we confirm that the NMSE is lower in the single-cycle dynamic state. Therefore, we choose k_{inj} to be 0.05 where the PRC has optimal consistency. (ii) Figures 4(b1) and 4(b2) illustrate the influence of frequency detuning $\Delta\nu$ on the reservoir of the PRC. From Fig. 4(b1), we observe that when $\Delta\nu$ changes in the region from -20 GHz to 20 GHz, the system undergoes a transition from a single-cycle state to chaos and then returns to another single-cycle state. Figure 4 (b2) is the corresponding NMSE with the changing of $\Delta\nu$. We find that a smaller NMSE can be obtained at the edge of chaos. Thus, -5 GHz frequency detuning is set in the PRC, where the input signal can be mapped to the high-dimensional space and thus achieves the great prediction performance. (iii) we discuss the effects of k_f when k_{inj} is 0.05 and $\Delta\nu$ is -5 GHz. It can be seen from the bifurcation diagram in Fig. 4(c1) that the system quickly enters a chaotic state after a short single-cycle state with the increase of k_f. The trend of NMSE in Fig. 4 (c2) is that the NMSE gradually decreases and then increases with increasing k_f. The minimum NMSE can be obtained when k_f is 0.04. This is because the feedback strength k_f reflects the short-term memory ability of the reservoir [17]. Meanwhile, it also should be considered that an excessive feedback strength takes the system into an extremely complex dynamic state. At this point, the transient response will be disturbed easily and then the consistency of the system will be degraded [14, 17]. Considering the balance between the consistency and short-term memory, we thus choose a feedback strength k_f of 0.04.

After the above training processes, the meta-parameters of PRC have been adjusted to their optimum values, respectively. Under these conditions, we utilize the trained single-node RPC to forecast the future chaotic behaviors of the ECSL. Figure 5(a) shows the numerically produced chaotic time series from the ECSL, while Figure 5(b) is the forecasting chaotic time series by our PRC. Comparing them, we can confirm that the upcoming chaotic time series can be continuously predicted by the single-node PRC in a time duration over 2 ns. The associated NMSE is calculated to be less than 0.1.

Furthermore, we numerically produce new chaotic time series from another ECSL with different internal parameters from those used in the R-L to investigate the flexibility of our PRC. Note that, the feedback delay



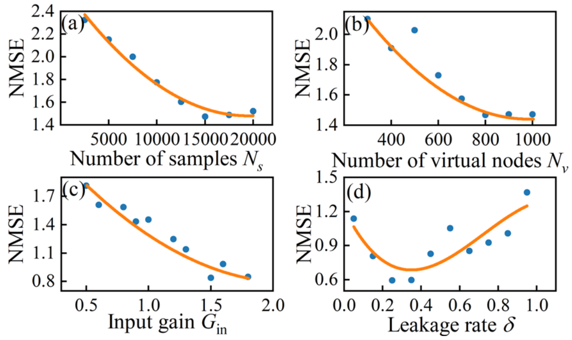


Fig. 5. (a) Simulated chaotic intensity time series of the ECSL with a 1 μs feedback delay time. (b) Predicted time series of the PRC.

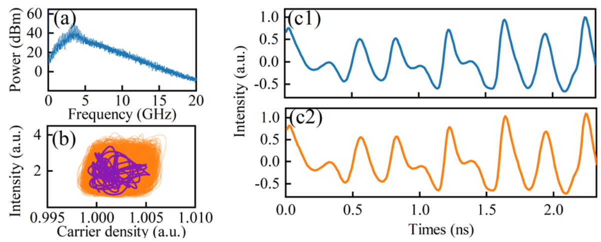


Fig. 6. (a) Power spectrum, (b) Phase diagram and (c1) Simulated chaotic intensity time series of the ECSL with different parameters from those used in the R-L for PRC. (c2) Forecasted time series of the PRC. The orange and the violet attractors are plotted for the 500 ns and the 5 ns chaotic time series, respectively.

time in this ECSL here is set to be a more commonly used value of a few nanoseconds. Quantitatively, these associated parameters and their values are as follows [29]: photon lifetime $\tau_p = 2$ ps, carrier lifetime $\tau_s = 2$ ns, carrier density at transparency $N_0 = 3\times10^{24}$ m^{-3}, gain saturation coefficient $\varepsilon = 2.5\times10^{-23}$ and feedback delay time $\tau = 5$ ns. Figures 6(a) and (b) show a typical power spectrum of the simulated chaotic time series and its associated chaotic attractor. After re-training the PRC with these chaotic time series, we find that the future chaotic behaviors of this new ECSL can be forecasted well, as shown in Fig. 6(c). Comparing the numerical simulated chaotic time series [Fig. 6(c1)] with that forecasted by our PRC, we confirm that the prediction time is about 2.3 ns with a NMSE lower than 0.1. In contrast with Fig. 5, it can be determined that the prediction performance of the PRC is not greatly affected, although chaos comes from two different ECSLs.

We point out that such good prediction results benefit from the unique short-term memory processed by the PRC. As known, whether a network has a memory property is critical for processing time-dependent signals. Due to the delay loop, the past information of the input signal can mix with the current input in the reservoir layer so that our PRC has an inherent memory property. As the input of the reservoir in our prediction task, optical chaos is a kind of signal with a short-time relevance. Therefore, the memory in our PRC needs to clear after some time to allow responses to be influenced only by the recent past. As demonstrated by us in [30], this requirement can be satisfied when the reservoir operates (in the absence of input) at the edge of the unstable region by optimizing its meta-parameters such as injection strength, feedback strength, and frequency detuning.

In conclusion, we have numerically demonstrated a method based on single-node PRC to continuously forecast the chaotic behaviors of ECSLs. Moreover, we have investigated the effect of some critical hyperparameters (i.e., virtual node number, the training data size, input gain, leakage rate, injection strength, frequency detuning, feedback strength) on prediction performance. Final results demonstrate that using our method, the upcoming chaotic time series can be continuously predicted in a time duration over 2 ns with a normalized mean squared error lower than 0.1. The forecasting performance may be further improved by introducing parallelizing techniques. Considering the good performance and simple optical devices, we believe that the single-node PRC offers a hardware-friendly approach for complex chaos prediction.

Funding. National Natural Science Foundation of China under Grants (62175177, U19A2076, and 61927811); Program for Guangdong Introducing Innovative and Entrepreneurial Teams; Natural Science Foundation of Shanxi Province under Grant (201901D211116; 201901D211077).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this Letter are not publicly available at this time but be obtained from the authors upon reasonable request.

REFERENCES

1. N. D. Truong, J. Y. Haw, S. M. Assad, P. K. Lam, and O. Kavehei, IEEE T. Inf. Foren. Sec. **14**, 403 (2018).
2. N. Lv, T. Chen, S. Zhu, J. Yang, Y. Ma, J. Jing, and J. Lin, Secur. Commu. Netw. **2020**, 4241713 (2020).
3. Y. Zhang, Y. Xu, and Z. Wang, Chaos Solitons Fractals **42**, 1003-1009 (2009).
4. J. Ke, L. Yi, Z. Yang, Y. Yang, Q. Zhuge, Y. Chen, and W. Hu, Opt. Lett. **44**, 5776 (2019).
5. G. Shmueli, Stat. Sci. **25**, 289 (2010).
6. K. Chakraborty, K. Mehrotra, C. K. Mohan, and S. Ranka, Neural Netw. **5**, 961 (1992).
7. D. J. Gauthier, E. Bollt, A. Griffith, and W. A. Barbosa, "Nat. Commun. **12**, 1 (2021).
8. Y. Zhong, J. Tang, X. Li, B. Gao, H. Qian, and H. Wu, Nat. Commun. **12**, 1 (2021).
9. A. M. Logar, E. M. Corwin, and W. J. Oldham, IEEE Int. Conf. Neural Networks, 1129 (1993).
10. J.-S. Zhang and X.-C. Xiao, Chinese Phys. Lett. **17**, 88 (2000).
11. X. Z. Li, B. Sheng, and M. Zhang, Opt. Lett. **47**, 2822 (2022).
12. J. Pathak, B. Hunt, M. Girvan, Z. Lu, and E. Ott, Phys. Rev. Lett. **120**, 024102 (2018).
13. P. Amil, M. C. Soriano, and C. Masoller, Chaos **29**, 113111 (2019).
14. J. Bueno, D. Brunner, M. C. Soriano, and I. Fischer, Opt. Express **25**, 2401 (2017).
15. Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, Sci. Rep. **2**, 1 (2012).
16. L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutiérrez, L. Pesquera, C. R. Mirasso, and I. Fischer, Opt. Express **20**, 3241 (2012).
17. L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, Phys. Rev. X **7**, 011015 (2017).
18. F. Duport, A. Smerieri, A. Akrout, M. Haelterman, and S. Massar, Sci. Rep. **6**, 1 (2016).
19. F. Duport, A. Smerieri, A. Akrout, M. Haelterman, and S. Massar, J. Lightwave Technol. **34**, 2085 (2016).
20. Y. Kawaguchi, T. Okuma, K. Kanno, and A. Uchida, Opt. Express **29**, 2442 (2021).
21. M. C. Soriano, J. García-Ojalvo, C. R. Mirasso, and I. Fischer, Rev. Mod. Phys. **85**, 421 (2013).
22. P. Li, Y.-C. Wang, A.-B. Wang, L.-Z. Yang, M.-J. Zhang, and J.-Z. Zhang, Opt. Express **20**, 4297 (2012).
23. L. Jiang, J. Feng, L. Yan, A. Yi, S.-S. Li, H. Yang, Y. Dong, L. Wang, A. Wang, and Y. Wang, Opt. Lett. **47**, 2382 (2022).
24. M. Zhang, A. Wang, and Y. Wang, in 14th International Conference on Optical Communications and Networks (ICOCN), (IEEE, 2015).
25. Y. Wang, B. Wang, and A. Wang, IEEE Photonic. Tech. L. **20**, 1636 (2008).
26. T. Heil, I. Fischer, W. Elsässer, J. Mulet, and C. R. Mirasso, Phys. Rev. Lett. **86**, 795 (2001).

27. A. E. Hoerl and R. W. Kannard, Dent. Tech. **12**, 55 (1970).
28. P. Händel, IEEE Microw. Wireless Compon. Lett. **28**, 1047 (2018).
29. D. Rontani, A. Locquet, M. Sciamanna, D. S. Citrin, and S. Ortin, IEEE J. Quantum Electron. **45**, 879 (2009).
30. Q. Cai, Y. Guo, P. Li, A. Bogris, K. A. Shore, Y. Zhang, Y. Wang, Photon. Res. **9**, B1(2021).

FULL REFERENCES

1. N. D. Truong, J. Y. Haw, S. M. Assad, P. K. Lam, and O. Kavehei, "Machine learning cryptanalysis of a quantum random number generator," IEEE T. Inf. Foren. Sec. **14**, 403-414 (2018).
2. N. Lv, T. Chen, S. Zhu, J. Yang, Y. Ma, J. Jing, and J. Lin, "High-efficiency min-entropy estimation based on neural network for random number generators," Secur. Commu. Netw. **2020**, 4241713 (2020).
3. Y. Zhang, Y. Xu, and Z. Wang, "GM (1, 1) grey prediction of Lorenz chaotic system," Chaos Solitons Fractals **42**, 1003-1009 (2009).
4. J. Ke, L. Yi, Z. Yang, Y. Yang, Q. Zhuge, Y. Chen, and W. Hu, "32 Gb/s chaotic optical communications by deep-learning-based chaos synchronization," Opt. Lett. **44**, 5776-5779 (2019).
5. G. Shmueli, "To explain or to predict?" Stat. Sci. **25**, 289-310 (2010).
6. K. Chakraborty, K. Mehrotra, C. K. Mohan, and S. Ranka, "Forecasting the behavior of multivariate time series using neural networks," Neural Netw. **5**, 961-970 (1992).
7. D. J. Gauthier, E. Bollt, A. Griffith, and W. A. Barbosa, "Next generation reservoir computing," Nat. Commun. **12**, 1-8 (2021).
8. Y. Zhong, J. Tang, X. Li, B. Gao, H. Qian, and H. Wu, "Dynamic memristor-based reservoir computing for high-efficiency temporal signal processing," Nat. Commun. **12**, 1-9 (2021).
9. A. M. Logar, E. M. Corwin, and W. J. Oldham, "A comparison of recurrent neural network learning algorithms," IEEE Int. Conf. Neural Networks, 1129-1134 (1993).
10. J.-S. Zhang and X.-C. Xiao, "Predicting chaotic time series using recurrent neural network," Chinese Phys. Lett. **17**, 88 (2000).
11. X. Z. Li, B. Sheng, and M. Zhang, "Predicting the dynamical behaviors for chaotic semiconductor lasers by reservoir computing," Opt. Lett. **47**, 2822-2825 (2022).
12. J. Pathak, B. Hunt, M. Girvan, Z. Lu, and E. Ott, "Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach," Phys. Rev. Lett. **120**, 024102 (2018).
13. P. Amil, M. C. Soriano, and C. Masoller, "Machine learning algorithms for predicting the amplitude of chaotic laser pulses," Chaos **29**, 113111 (2019).
14. J. Bueno, D. Brunner, M. C. Soriano, and I. Fischer, "Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback," Opt. Express **25**, 2401-2412 (2017).
15. Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, "Optoelectronic reservoir computing," Sci. Rep. **2**, 1-6 (2012).
16. L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutiérrez, L. Pesquera, C. R. Mirasso, and I. Fischer, "Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing," Opt. Express **20**, 3241-3249 (2012).
17. L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, "High-speed photonic reservoir computing using a time-delay-based architecture: Million words per second classification," Phys. Rev. X **7**, 011015 (2017).
18. F. Duport, A. Smerieri, A. Akrout, M. Haelterman, and S. Massar, "Fully analogue photonic reservoir computer," Sci. Rep. **6**, 1-12 (2016).
19. F. Duport, A. Smerieri, A. Akrout, M. Haelterman, and S. Massar, "Virtualization of a photonic reservoir computer," J. Lightwave Technol. **34**, 2085-2091 (2016).
20. Y. Kawaguchi, T. Okuma, K. Kanno, and A. Uchida, "Entropy rate of chaos in an optically injected semiconductor laser for physical random number generation," Opt. Express 29, 2442-2457 (2021).
21. M. C. Soriano, J. García-Ojalvo, C. R. Mirasso, and I. Fischer, "Complex photonics: Dynamics and applications of delay-coupled semiconductors lasers," Rev. Mod. Phys. **85**, 421 (2013).
22. P. Li, Y.-C. Wang, A.-B. Wang, L.-Z. Yang, M.-J. Zhang, and J.-Z. Zhang, "Direct generation of all-optical random numbers from optical pulse amplitude chaos," Opt. Express **20**, 4297-4308 (2012).
23. L. Jiang, J. Feng, L. Yan, A. Yi, S.-S. Li, H. Yang, Y. Dong, L. Wang, A. Wang, and Y. Wang, "Chaotic optical communications at 56 Gbit/s over 100-km fiber transmission based on a chaos generation model driven by long short-term memory networks," Opt. Lett. **47**, 2382-2385 (2022).
24. M. Zhang, A. Wang, and Y. Wang, "Chaotic light generation and its application in communication, sensor, radar," in 14th International Conference on Optical Communications and Networks (ICOCN), (IEEE, 2015), 1-3.
25. Y. Wang, B. Wang, and A. Wang, "Chaotic correlation optical time domain reflectometer utilizing laser diode," IEEE Photonic. Tech. L. **20**, 1636-1638 (2008).
26. T. Heil, I. Fischer, W. Elsässer, J. Mulet, and C. R. Mirasso, "Chaos synchronization and spontaneous symmetry-breaking in symmetrically delay-coupled semiconductor lasers," Phys. Rev. Lett. **86**, 795 (2001).
27. A. E. Hoerl and R. W. Kannard, "Ridge regression: Biased estimation for nonorthogonal problems," Dent. Tech. **12**, 55-67 (1970).
28. P. Händel, "Understanding normalized mean squared error in power amplifier linearization," IEEE Microw. Wireless Compon. Lett. **28**, 1047-1049 (2018).
29. D. Rontani, A. Locquet, M. Sciamanna, D. S. Citrin, and S. Ortin, "Time-delay identification in a chaotic semiconductor laser with optical feedback: a dynamical point of view," IEEE J. Quantum Electron. **45**, 879-1891 (2009).
30. Q. Cai, Y. Guo, P. Li, A. Bogris, K. A. Shore, Y. Zhang, Y. Wang, "Modulation format identification in fiber communications using single dynamical node-based photonic reservoir computing," Photon. Res. **9**, B1(2021).