Machine learning based on reservoir computing with time-delayed optoelectronic and photonic systems

Yanne K. Chembo

Submitted: 21 July 2019. Accepted: 26 November 2019. Published Online: 08 January 2020
Machine learning based on reservoir computing with time-delayed optoelectronic and photonic systems

Cite as: Chaos 30, 013111 (2020); doi: 10.1063/1.5120788
Submitted: 21 July 2019 · Accepted: 26 November 2019 · Published Online: 8 January 2020

Yanne K. Chembo

AFFILIATIONS
Department of Electrical and Computer Engineering, Institute for Research in Electronics and Applied Physics (IREAP), University of Maryland, 8279 Paint Branch Dr., College Park, Maryland 20742, USA

Note: This paper is part of the Focus Issue, "When Machine Learning Meets Complex Systems: Networks, Chaos and Nonlinear Dynamics."

Electronic mail: ykchembo@umd.edu

ABSTRACT
The concept of reservoir computing emerged from a specific machine learning paradigm characterized by a three-layered architecture (input, reservoir, and output), where only the output layer is trained and optimized for a particular task. In recent years, this approach has been successfully implemented using various hardware platforms based on optoelectronic and photonic systems with time-delayed feedback. In this review, we provide a survey of the latest advances in this field, with some perspectives related to the relationship between reservoir computing, nonlinear dynamics, and network theory.

Machine learning (ML) based on neural networks has achieved a deep and transformational impact on modern science and technology. This bioinspired computational approach, sometimes referred to as neuromorphic computing, has emerged as one of the most powerful paradigms to solve certain classes of problems where Turing-Von Neumann machines are computationally inefficient. ML has been astounding successful in certain contexts such as data mining, sorting, or classification, thereby proving its superiority in pattern recognition tasks performed in parameter spaces too large to be searched using conventional—and exponentially time-consuming—algorithms. However, beyond mere pattern recognition, ML has succeeded as well to achieve major breakthroughs in areas where artificial intelligence was not expected to bridge the gap with human intelligence (such as for the games of Go or Chess, for example). The main upcoming challenge in ML is to achieve a deeper understanding of the core mechanisms that ensure high performance in these bioinspired computers.

I. INTRODUCTION
The operating principle of ML using neural networks is based on two main phases, namely, training and testing. In a nutshell, a known dataset is used to train the network and to optimize its connectivity for that particular set of data. Unknown information is then fed to the network, which, therefore, processes the incoming signals using the previously optimized coupling coefficients (or weights). However, this network optimization can be quite energy- and time-inefficient depending on many factors such as the complexity of the task, the size of the network, the nature of the nonlinear nodes (physical benchmark), the coupling architecture between the nodes, and the operational bandwidth of connectivity links.

Most of these difficulties can be overcome using the idea of reservoir computing. This approach finds its roots in the concepts of a liquid-state machine (LSM, see Maass et al.) and one of echo state networks (ESNs, see Jaeger and Haas) that were later on unified under the concept of reservoir computing (or RC, see Verstraeten et al.). As shown in Fig. 1, reservoir computers consist of three layers, labeled as input, reservoir, and output. The most salient feature of reservoir computing is that only the readout coefficients of the output layer have to be trained and optimized, while the input and reservoir layers are fixed (see Lukosevicius and Jaeger). This radical simplification in the network’s architecture was ideally suited for hardware implementation in several physical platforms, as reviewed by Tanaka et al. It has also been a key factor for the implementation of reservoir computing using time-delayed dynamical systems.
Since the pioneering work of Arecchi et al., it is well known that time-delayed systems are infinite-dimensional, and, as a consequence, they can be mapped into a spatiotemporal representation (see also the recent review by Yanchuk and Giacomelli). Consequently, a large number of virtual neurons can be in principle excited in time-delayed systems, exactly as they would in a spatially extended system.

Appeltant et al. used a Mackey-Glass nonlinear electronic circuit to propose the first implementation of reservoir computing based on a time-delayed nonlinear dynamics. Their reservoir computing architecture involved a single nonlinear node $f_{NL}$ that was combined with a delayed feedback loop of round-trip time $T$. A total number $N$ of virtual neurons are emulated by multiplying the input signal to be processed by a time-domain mask $m(t)$, which is a $T$-periodic and multilevel step function with segments of equal duration $\delta T = T/N$.

Up to several thousand virtual neurons can thereby be activated, in order to process the incoming information in the highly dimensional reservoir, as shown in Fig. 2 where the generic architecture of input, reservoir, and output layers in time-delay-based reservoir computers is presented.

This novel approach to reservoir computing was later on successfully translated to optoelectronic and photonic platforms (see, for example, review articles by Van der Sande et al. and Brunner et al.). The suitability of time-delayed reservoir computing architectures for optoelectronic and photonic systems is justified by several reasons, the main one being that in optical systems, addressing physical nodes in a reservoir and sequentially adjusting their coupling weights is quite a complex task, while it is a much simpler endeavor using virtual nodes embedded in a temporal delay line. As far as optoelectronic systems are concerned, they are characterized by a spectral duality that allows them to process input signals either in the microwave domain (radar signals, wireless communications) or in the light wave domain (optical fiber data, video signals, etc.). They also feature a very broad bandwidth (up to 100 GHz) and are, therefore, capable of ultrafast information processing.

A deeper understanding of the core mechanisms permitting reservoir computing using these time-delayed systems is still nevertheless required in order to take full advantage of their computing power. Despite being underdeveloped at this date, research on this important aspect of reservoir computing could benefit from various techniques borrowed from the theory of nonlinear and complex dynamical systems.
This article is organized as follows. We discuss in Sec. II the main formalisms that have been developed to understand and analyze reservoir computing. Section III provides a short overview of the various implementations of reservoir computing systems based on time-delayed optoelectronic oscillators, while Sec. IV proposes a survey or the research related to all-optical RC. Some perspectives related to future challenges with optoelectronic and photonic reservoir computing are explored in Sec. V, with an emphasis on real-time and energy-efficient operation. We will consider as well other aspects such as cross-fertilization with some fundamental concepts associated with the nonlinear dynamics of continuous spatiotemporal systems and discrete networks. The last section concludes the article.

II. NONLINEAR DYNAMICS OF RC BASED ON TIME-DELAYED SYSTEMS

In conventional reservoir computing architectures (see Fig. 1), the internal state \( x(n) \) of the reservoir and the output signal \( y(n) \) are updated at each discrete time step \( n \) following

\[
x(n) = f_{NL}[W_{in} \cdot u(n) + W \cdot x(n - 1)],
\]

\[
y(n) = W_{out} \cdot x(n),
\]

where \( f_{NL} \) is a vector nonlinear function (equivalent to the same scalar function \( f_{NL} \) for each row) and \( u(n) \) is the input signal that is injected in the reservoir. In the case of supervised learning, the optimal readout matrix \( W_{out} \) is obtained by ridge regression following

\[
W_{out}^{\text{opt}} = M_{x} \cdot M_{x}^T \cdot [M_{x} \cdot M_{x}^T + \lambda \cdot I]^{-1},
\]

where \( M_{x} \) is a suitably designed matrix that concatenates the internal state \( x \) obtained with some training input vectors \( u \), \( M_{x} \) is the target matrix that yields the desired classification outcome, \( I \) is the identity matrix, and \( \lambda \ll 1 \) is a small regularization factor required to circumvent the ill-posedness of the inversion problem. Indeed, finding the optimal readout coefficients via a one-step regression procedure in reservoir computing provides significantly faster results than most ML architectures where all the connectivity matrices have to be optimized via a multistep, slowly converging procedure.

Unlike in the conventional reservoir computing architectures that are fed by multidimensional and discrete inputs \( u(n) \), reservoir computing systems based on time-delay are excited with one-dimensional and continuous input signals \( u(t) \). Both approaches become structurally equivalent when the time-domain signal \( u(t) \) is spatiotemporally mapped into the reservoir via a \( T \)-periodic temporal mask \( m(t) \) that plays the role of the input connectivity matrix \( W_{in} \).

An important specificity of reservoir computing based on a nonlinear delayed feedback loop is that information is injected and retrieved from the reservoir via time-multiplexing in order to address the \( N \) virtual nodes with \( \delta T = T/N \) temporal spacing. Both experiments and simulations show that optimal RC efficiency is achieved when the oscillators are biased close to the threshold, and it is the information signal \( u(t) \) that drives the dynamics of the system, which responds with a high-dimensional temporal dynamics unfolding in an infinite-dimensional state space. Hence, optimizing the readout matrix \( W_{out} \) can be interpreted finding the optimal coefficients of separation hyperplanes that permit pattern classification.

III. OPTOELECTRONIC ARCHITECTURES OF RC

Optoelectronic oscillators (OEOs) are autonomous systems characterized by a feedback loop where the signal alternatively circulates in the optical and electrical forms, with a round-trip duration corresponding to the time delay.

A particularly important class of OEOs is based on a Ikeda-like architecture that can be characterized by four main elements, namely, a nonlinear transfer function \( f_{NL} \), a delay line \( T \), an amplifier \( \beta \), and a linear filter \( H \) with an impulse function \( h(t) \). Typically, these OEOs have one or several fixed points when \( \beta \) is low, and when the gain is continuously increased, periodic oscillations are triggered via a Hopf bifurcation beyond a critical value \( \beta_{cr} \). Further increase of the gain leads to multiperiodic and then hyperchaotic oscillations. In recent years, OEOs have been investigated thoroughly from the standpoint of nonlinear dynamics. Optoelectronic oscillators have also found numerous applications related to future challenges with optoelectronic and photonic reservoir computing. Some perspectives related to future challenges with optoelectronic and photonic reservoir computing are explored in Sec. V

FIG. 3. An example of OEO-based reservoir computing architecture. The dynam-ical variable in the feedback loop is \( x(t) \propto V_{RF}(t) \), corresponding to the oscillation at the radiofrequency input of the electro-optic Mach-Zehnder modulator. PC: polar-ization controller; MZM: Mach-Zehnder modulator; DL: delay line; PD: photodiode; Amp: RF amplifier; and MC: microwave coupler.
concept of OEO-based reservoir computing has been extended by other research groups using various modifications to the core architecture. The computing efficiency of OEO-based RC has been tested with a wide variety of tasks that can be organized into two main categories: classification and prediction problems.

As far as pattern recognition problems are concerned, several approaches have been proposed to demonstrate high performance to classify audio signals, images, or temporal laser signals. A reservoir computer based on an OEO oscillating in wavelength was introduced by Martinenghi et al.\textsuperscript{10} for spoken-digit recognition, and it involved a field programmable gate array (FPGA) board that was used to expand the reservoir size by emulating multiple delay lines. Hermans et al.\textsuperscript{11} proposed a backpropagation algorithm for the TIMIT benchmark, which is an acoustic-phonetic continuous speech test. This protocol, which is similar to the standard procedure in multilayer perceptrons, was shown to provide superior performances comparatively to the regressive reservoir computing algorithms. A major breakthrough was achieved later on by Larger et al.\textsuperscript{12} as they used a differential phase shift keying (DPSK) element as a nonlinear phase-to-intensity converter in the OEO loop to demonstrate the state-of-the-art speed performance of one-million words per second for two spoken-digit tests, namely, TI-46 and AURORA-2. Beyond audio signals, OEOs have also been considered by Jin et al.\textsuperscript{13} to perform numerical simulations of reservoir computing for handwritten numeral recognition with the MNIST database. In the optical domain, one of the main applications of OEO-based reservoir computing is optical header recognition, as investigated by Qin et al.,\textsuperscript{14} Zhao et al.,\textsuperscript{15} and Bao et al.\textsuperscript{16}

OEO-based reservoir computing was shown as well to be efficient for time-series prediction tasks. For example, Soriano et al.\textsuperscript{17} have demonstrated that multilevel (instead of the standard two-level) masks lead to higher robustness to noise, while Tezuka et al.\textsuperscript{18} evidenced the relevance of mutually coupling OEOs in order to reduce the mask modulation frequency. Other architecture innovations

### TABLE I. Some of the performances experimentally achieved by OEO-based reservoir computers, with a specification of the number of virtual nodes used in the computation. FPGA: field programmable gate array; NMSE: normalized mean-square error; NRMSE: normalized root-mean-square error; SER: symbol error rate; and WER: word error rate.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Optoelectronic RC architecture</th>
<th>Nodes</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antonik et al.\textsuperscript{19}</td>
<td>Bandpass OEO with voltage dynamical variable; FPGA-enabled online training</td>
<td>50</td>
<td>Nonlinear channel equalization: SER = 6 × 10\textsuperscript{-4} with SNR = 32 dB</td>
</tr>
<tr>
<td>Duport et al.\textsuperscript{20}</td>
<td>Bandpass OEO with voltage dynamical variable; simultaneous RC for several tasks</td>
<td>50</td>
<td>(i) NARMA10: NMSE = 0.18; (ii) nonlinear channel equalization: SER ∼ 10\textsuperscript{-4} with SNR = 32 dB; (iii) signal radar forecasting: NMSE ∼ 10\textsuperscript{-3}</td>
</tr>
<tr>
<td>Hermans et al.\textsuperscript{21}</td>
<td>Bandpass OEO with voltage dynamical variable; optimization via backpropagation</td>
<td>(i) 80; (ii) 200</td>
<td>(i) NARMA10: NRMSE = 0.185; (ii) acoustic-phonetic continuous speech corpus TIMIT: WER = 35%</td>
</tr>
<tr>
<td>Larger et al.\textsuperscript{22}</td>
<td>Low-pass OEO with voltage dynamical variable; proof-of-concept experiment</td>
<td>400</td>
<td>(i) Spoken-digit recognition TI46: WER = 0.04%; (ii) Santa Fe time-series prediction: NMSE ∼ 12%</td>
</tr>
<tr>
<td>Larger et al.\textsuperscript{23}</td>
<td>Bandpass OEO with optical phase dynamical variable; record of one-million words per second for spoken-digit recognition tasks</td>
<td>1113</td>
<td>(i) Spoken-digit recognition AURORA-2: WER = 6%; (ii) spoken-digit recognition TI46: WER = 0.04%</td>
</tr>
<tr>
<td>Martinenghi et al.\textsuperscript{24}</td>
<td>Low-pass OEO with wavelength dynamical variable; up to 15 delays implemented via FPGA</td>
<td>150</td>
<td>(i) Spoken-digit recognition TI46: WER = 0.6%</td>
</tr>
<tr>
<td>Ortín et al.\textsuperscript{25}</td>
<td>Low-pass OEO with optical phase dynamical variable; unification of reservoir computing and extreme learning machines</td>
<td>400</td>
<td>(i) Cancer classification with microarray data: &lt;5% error (down to ~2% with &gt;1000 nodes); (ii) nonlinear channel equalization: SER ∼ 10\textsuperscript{-3} with SNR = 28 dB</td>
</tr>
<tr>
<td>Paquot et al.\textsuperscript{26}</td>
<td>Bandpass OEO with voltage dynamical variable; proof-of-concept experiment</td>
<td>(i) 50; (ii) 50; (iii) 200</td>
<td>(i) NARMA10: NRMSE = 0.17; (ii) nonlinear channel equalization: SER = 1.3 × 10\textsuperscript{-4} with SNR = 28 dB; (iii) spoken-digit recognition TI46: WER = 0.4%</td>
</tr>
<tr>
<td>Qin et al.\textsuperscript{27}</td>
<td>Bandpass OEO with voltage dynamical variable; optical packet header recognition from 3 to 32 bits</td>
<td>400</td>
<td>Optical packet header recognition: SER = 1.25%</td>
</tr>
<tr>
<td>Soriano et al.\textsuperscript{28}</td>
<td>Bandpass OEO with voltage dynamical variable; multilevel preprocessing mask</td>
<td>400</td>
<td>Santa Fe time-series prediction: NMSE ∼ 2%</td>
</tr>
</tbody>
</table>
include unification with extreme learning machines,\textsuperscript{21} time-domain multiplexing within the same reservoir for the simultaneous processing of several tasks,\textsuperscript{11} computation with fully analog input and output layers,\textsuperscript{16} recurrent reservoir computing schemes where the output signals are reinjected into the reservoir,\textsuperscript{21} or multiloop configurations.\textsuperscript{22}

We refer the reader to Table I for an overview of some experimental results obtained with OEO-based RC.

IV. ALL-OPTICAL ARCHITECTURES OF RC

All-optical systems with time-delayed feedback provide an ideal platform to investigate complex dynamics in a wide range of timescales (see the review article by Soriano et al.\textsuperscript{89}). The applications have been numerous as well, including optical communications,\textsuperscript{55,56} random number generation,\textsuperscript{86,90} and various subfields of microwave photonics.\textsuperscript{89} Several architectures have been proposed in the literature to implement all-optical RC, such as the one displayed in Fig. 4.

One of the earliest all-optical RC configurations was introduced by Duport et al.,\textsuperscript{39} with a system where the nonlinear node was a semiconductor optical amplifier (SOA). Soon after, Dejongckheere et al.\textsuperscript{38} proposed a RC scheme where the nonlinearity was originating from a semiconductor saturable absorber mirror (SESAM). In both cases, the RC performance was theoretically and experimentally investigated with various tasks such as nonlinear channel equalization, the IPIX radar test (sea clutter radar data), and the spoken-digit recognition test T46.

Another popular framework for all-optical RC is based on using an external-cavity semiconductor laser as a nonlinear node. This approach was introduced by Brunner et al.,\textsuperscript{37} in a pioneering experiment where they demonstrated chaotic time-series, spoken-digit, and speaker recognition at data rates beyond 1 Gbyte/s. Further analysis was shortly after provided by Hicke et al.\textsuperscript{39} with regard to the performance of this system.

The efficiency of this ECSL-based RC architecture is intimately associated with the concept of consistency, initially introduced by Uchida et al.\textsuperscript{90} Along that line, Nakayama et al.\textsuperscript{89} studied the relationship between the consistency of the laser dynamics and RC performance. They compared the RC performance depending on the analog, digital, or chaotic nature of the input temporal mask and showed that the chaotic mask improved the efficiency of the reservoir owing to the complexity of its dynamical response. Further research by Kuriki et al.\textsuperscript{91} confirmed the relevance of such chaotic masks for RC. On the other hand, Bueno et al.\textsuperscript{89} investigated the optimal conditions ECSL-based RC by evidencing the interplay between consistency and memory, thereby permitting the identification of the most suitable parameters for various prediction tasks.

As far as optical communications are concerned, the potential of ECSL-based RC has been initially considered for optical packet header recognition.\textsuperscript{22} More recently, other tasks such as postprocessing for fiber distortion correction and fiber transmission equalization have been successfully demonstrated.\textsuperscript{94,95}

Alternative architectures of ECSL-based RC have recently been explored in the literature, such as semiconductor lasers with double optical feedback and optical injection,\textsuperscript{10} mutually delay-coupled semiconductor lasers,\textsuperscript{95} or architectures where the information is injected in the electrical electrode of the laser.\textsuperscript{90,96}

Vertical cavity surface-emitting lasers (VCSELs) have been proposed as well for the purpose of all-optical RC. These lasers have the specificity to display a polarization dynamics that provides additional degrees of freedom from the dynamical and computational standpoints. Vatin et al.\textsuperscript{94} proposed to implement RC using a VCSEL with time-delayed feedback and optical injection and theoretically demonstrated the technological relevance of this system. This concept was later on successfully implemented at the experimental level with benchmark tasks such as chaotic time-series prediction and nonlinear channel equalization.\textsuperscript{19} It was recently shown that VCSELs are suitable as well for multiplexing and parallel computation in configurations featuring mutual coupling,\textsuperscript{30} double optical feedback,\textsuperscript{22} or polarized optical feedback.\textsuperscript{106}

Other architectures for all-optical RC based on time-delayed feedback have been explored in the literature, characterized by distinctive features such as coherently driven passive cavities,\textsuperscript{36} semiconductor ring lasers,\textsuperscript{46,101} diode-pumped erbium-doped microchip lasers,\textsuperscript{81} photonic integrated circuits,\textsuperscript{92} optical neurons based on optical amplifiers,\textsuperscript{102} or semiconductor laser networks with optical feedback.\textsuperscript{103}

Finally, one should note that another important category of all-optical RC systems is based on photonic networks where the nodes are passive or active microrings.\textsuperscript{104-112}

We refer the reader to Table II for an overview of some experimental results achieved with all-optical RC.
TABLE II. Some of the performances experimentally achieved by all-optical reservoir computers, with a specification of the number of virtual nodes used in the computation. ECSL: external-cavity semiconductor laser; NMSE: normalized mean-square error; PAM-4: 4-level pulse amplitude modulation; SER: symbol error rate; VCSEL: vertical cavity surface-emitting laser; WER: word error rate.

<table>
<thead>
<tr>
<th>Reference</th>
<th>All-optical RC architecture</th>
<th>Nodes</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argyris et al.14</td>
<td>ECSL; bit pattern recognition</td>
<td>66</td>
<td>BER improvement by up to two orders of magnitude; extension of the communication range by over 75%</td>
</tr>
<tr>
<td>Argyris et al.14</td>
<td>ECSL; distortion correction of optical fiber transmission lines via RC postprocessing</td>
<td>32</td>
<td>21- and 4.6-km transmission at 56 and 112 Gb/s, respectively, for 14 dBm launched peak power of a PAM-4 signal</td>
</tr>
<tr>
<td>Brunner et al.15</td>
<td>ECSL; Santa Fe chaotic time-series, spoken-digit and speaker recognition at data rates beyond 1 Gbyte/s</td>
<td>388</td>
<td>Chaotic time-series prediction with error &lt;10%; WER &lt; 1% for the spoken-digit test; error rate &lt;1% for speaker recognition</td>
</tr>
<tr>
<td>Bueno et al.16</td>
<td>ECSL; Mackey-Glass chaotic time-series prediction</td>
<td>330</td>
<td>NMSE ~ 5%</td>
</tr>
<tr>
<td>Dejonckheere et al.17</td>
<td>Semiconductor saturable absorber mirror (SESAM); nonlinear channel equalization, IPIX radar test, and spoken-digit recognition test T146</td>
<td>(i) 50; (ii) 50; (iii) 200</td>
<td>(i) Nonlinear channel equalization: SER &lt; 10^{-4} with SNR = 32 dB; (ii) IPIX radar: NMSE ~ 0.2%; (iii) spoken-digit recognition T146: WER = 2.6%</td>
</tr>
<tr>
<td>Duport et al.18</td>
<td>Semiconductor optical amplifier (SOA); nonlinear channel equalization, IPIX radar test, and spoken-digit recognition test T146</td>
<td>(i) 50; (ii) 50; (iii) 200</td>
<td>(i) Nonlinear channel equalization: SER &lt; 10^{-3} with SNR = 32 dB; (ii) IPIX radar: NMSE ~ 0.1%; (iii) spoken-digit recognition T146: WER = 3%</td>
</tr>
<tr>
<td>Hicke et al.19</td>
<td>ECSL; Santa Fe chaotic time-series, spoken-digit recognition</td>
<td>388</td>
<td>Chaotic time-series prediction: NMSE ~ 10%; WER = 0% for the spoken-digit test</td>
</tr>
<tr>
<td>Kuriki et al.20</td>
<td>ECSL with optical injection; analysis of the influence of the mask</td>
<td>177</td>
<td>NMSE ~ 10%</td>
</tr>
<tr>
<td>Nguimdo et al.21</td>
<td>Erbium-doped microchip laser with optical feedback; Santa Fe time-series prediction task</td>
<td>100</td>
<td>NMSE ~ 5%</td>
</tr>
<tr>
<td>Takano et al.22</td>
<td>Photonic integrated circuit; Santa Fe time-series prediction task</td>
<td>Up to 250</td>
<td>NMSE ~ 5%</td>
</tr>
<tr>
<td>Vatin et al.23</td>
<td>VCSEL with data insertion via optical injection; Santa Fe time-series prediction task and nonlinear channel equalization</td>
<td>390</td>
<td>(i) Santa Fe time-series prediction: NMSE = 1.6%; (ii) nonlinear channel equalization: SER = 1.5%</td>
</tr>
<tr>
<td>Vinckier et al.24</td>
<td>Coherently driven passive cavity; NARMA10, nonlinear channel equalization, and spoken-digit recognition test T146</td>
<td>(i) 50; (ii) 50; (iii) 200</td>
<td>(i) NARMA10: NMSE ~ 10%; (ii) nonlinear channel equalization: SER &lt; 10^{-4} with SNR = 32 dB; (iii) spoken-digit recognition T146: WER = 0% without noise and WER = 0.8% with noise at 3 dB SNR level.</td>
</tr>
<tr>
<td>Yue et al.25</td>
<td>ECSL with information injection by current modulation Santa Fe time-series prediction task</td>
<td>200–800</td>
<td>NMSE ~ 5%</td>
</tr>
</tbody>
</table>

V. PERSPECTIVES IN OPTOELECTRONIC AND PHOTONIC RC

Research on the topic of optoelectronic and photonic reservoir computing has already led to important milestones, as discussed in Sec. III. There are, however, several aspects in this area that could even more increase the impact of these research efforts in the near future.

Indeed, it is known that arrays of nonlinear coupled oscillators can display a wide variety of coherent behaviors,11–15 and it is expected that the capabilities of reservoir computing can be significantly expanded from the on-going research on networks of coupled OEOs.11–15 In general, networks in time-delayed OEOs are built either by physically coupling several oscillators,17–21 by emulating a complex network architecture using FPGA boards,22,23 or by using spatially coupled systems.24,25 Each of these configurations has the potential to provide a relevant extension to the existing architectures of OEO-based reservoir computers, as recently discussed by Hart et al.26 Recent works have also highlighted the
relevance of FPGA-based Boolean logic networks for both time-series prediction\textsuperscript{12} and pattern recognition,\textsuperscript{12} which could be associated with OEOs in order to gain the capability to handle optical signals.

As highlighted earlier, delay-dynamical systems are infinite-dimensional and, in that aspect, are analogous to spatially extended systems. Insightful theoretical results have been recently achieved in the context of spatiotemporal dynamics forecasting.\textsuperscript{12,13} In principle, these methodologies could be successfully translated in the realm of photonic reservoir computers, in combination with information-theoretic approaches\textsuperscript{10–12} and novel architectures.\textsuperscript{13}

One of the main challenges in reservoir computing is operation in real time. Indeed, off-line reservoir computing is suitable for certain tasks (e.g., data mining or classification), while many other applications require both training and testing to be performed in real time (e.g., data routing or monitoring). Achieving real-time reservoir computing with high efficiency implies solving critical technical problems, such as fast convergence to low-error rate operation during the online training phase, short processing latency (including pre- and postprocessing time), and high throughput bandwidth.

It, therefore, appears that reservoir computing has a core architecture that is inherently suitable for such real-time computing, because it is a machine learning paradigm where only the output layer needs to be trained, while the reservoir remains static, thereby greatly speeding the determination of the optimal weights. Another key advantage of OEO-based reservoir computers is that digital signal processing (DSP) boards can be inserted in the feedback loop in order to provide added functionalities, and FPGA boards rapidly emerged as particularly suitable platforms for this endeavor,\textsuperscript{10,12,13} In fact, it is logical to expect these FPGA-based OEOs to be competitive solutions for real-time reservoir computing, as shown by Antonik et al.,\textsuperscript{1} who implemented a gradient descent algorithm for the purpose of online training for nonlinear channel equalization. It should also be added that a fundamental property of reservoir computers, namely, fading-memory, is fully compatible with real-time operation as it requires the reservoir to progressively loose memory of remotely past states.

The constraint of energy-efficiency in reservoir computing should equally deserve much attention in the short term. Indeed, one of the main motivations in the area of bioinspired computing architectures is the fact that biological brains are outperforming artificial ones from the energetic standpoint. For example, the human brain merely needs a few tens of watts, while for certain tasks, supercomputers need a power consumption thousand or even million times higher for a similar or poorer result. In fact, this challenge of energy-efficiency was central in the pioneering work of Jaeger and Haas,\textsuperscript{5} which was focused on saving energy in wireless communication networks.

Finally, it is interesting to note that several concepts adjacent to photonic reservoir computing have been developed recently, such as OEO-based coherent Ising machines (or CIMs; see Bohm et al.\textsuperscript{12})\textsuperscript{12}. Cross-fertilization with other types of delay-based RC that involve physical principles different from those of photonics (see, for example, Torrejon et al.,\textsuperscript{10} Dion et al.,\textsuperscript{10} Estébanez et al.,\textsuperscript{17} and Riou et al.\textsuperscript{18}) would also provide further opportunity to develop heterogeneous platforms, which would favor the rapid and ubiquitous deployment of multiphysics machine learning hardware.

VI. CONCLUSION

In the last decade, photonic oscillators have been successfully used to perform reservoir computing for a large variety of tasks.

From the fundamental viewpoint, both the problems of optimal convergence and convergence rate are still open for reservoir computing, in general, and delay-based reservoir computing, in particular. As highlighted in Sec. V, cross-fertilization with closely related works in the domain of spatially extended systems and networks could offer new opportunities to advance our understanding of delay-based photonic reservoir computing.

The main attributes of photonic systems—particularly dual microwave/light wave operation, wide bandwidth, and compatibility with digital signal processing systems—place them at the interface between wireless and optical communication systems, with the potential to play an important role in the inclusion of reservoir computing in modern communication systems. Finally, addressing the usual constraints of size, weight, and power (SWAP), along with the main challenges of low-latency, high throughput, and real-time operation, is the essential route that will most likely allow OEO-based reservoir computers to have an impact beyond academic research.

ACKNOWLEDGMENTS

The author acknowledges financial support from the University of Maryland through the Minta Martin Fellowship.

REFERENCES

Chaos 30, 013111 (2020); doi: 10.1063/5.1207858
Published under license by AIP Publishing.


