OPEN

Deciphering Dynamical Nonlinearities in Short Time Series Using Recurrent Neural Networks

Radhakrishnan Nagarajan

Surrogate testing techniques have been used widely to investigate the presence of dynamical nonlinearities, an essential ingredient of deterministic chaotic processes. Traditional surrogate testing subscribes to statistical hypothesis testing and investigates potential differences in discriminant statistics between the given empirical sample and its surrogate counterparts. The choice and estimation of the discriminant statistics can be challenging across short time series. Also, conclusion based on a single empirical sample is an inherent limitation. The present study proposes a recurrent neural network classification framework that uses the raw time series obviating the need for discriminant statistic while accommodating multiple time series realizations for enhanced generalizability of the findings. The results are demonstrated on short time series with lengths ($L = 32, 64, 128$) from continuous and discrete dynamical systems in chaotic regimes, nonlinear transform of linearly correlated noise and experimental data. Accuracy of the classifier is shown to be markedly higher than $\gg 50\%$ for the processes in chaotic regimes whereas those of nonlinearly correlated noise were around ~50% similar to that of random guess from a one-sample binomial test. These results are promising and elucidate the usefulness of the proposed framework in identifying potential dynamical nonlinearities from short experimental time series.

Time series data can be realized by discretizing a continuous process in amplitude and time. Discretization in amplitude is a result of quantization whereas discretization in time can be achieved using an optimal sampling frequency (e.g. Nyquist rate)[1] for certain class of processes. Understanding the correlation structure is fundamental to time series analysis and can provide critical insights into its generative mechanism. On a related note, optimal parameters of a linearly correlated processes such as auto-regressive process can be estimated faithfully from their auto-correlation function (Yule-Walker equations)[1]. Auto-correlation in turn is related to their power-spectral density representing the distribution of the power across the various frequencies by the Wiener-Khintchine theorem[1]. Parametric as well as non-parametric approaches have been used widely for spectral estimation. Of interest is to note that non-parametric approaches such as subspace decomposition (Pisarenko Harmonic Decomposition)[1] estimate the dominant frequencies by eigen-decomposition of the corresponding Toeplitz matrix whose elements are essentially the auto-correlation function. On the other hand, correlation signatures in a given time series need not necessarily be linear. Nonlinear correlations can arise as a result of static nonlinearities as well as dynamical nonlinearities. Static nonlinearities are often attributed to the transfer function of a measurement device (e.g. sensor) that maps an analog or continuous process onto digital data. In contrast, dynamical nonlinearities such as those from nonlinear deterministic systems are a result of nonlinear coupling and can exhibit a wide-range of intricate behaviors including deterministic chaos[2–8]. Identifying chaos can be helpful in developing suitable approaches for their control[9,10]. Chaos has also been shown to have a wide-range of applications[11]. Break down in dynamical nonlinearities have also been shown to discriminate health and disease[3]. It is important to note that spectral analysis while useful for investigating narrow-band processes can be singularly unhelpful in adequately describing chaotic processes as they exhibit a broad-band spectrum similar to that of noise[12]. On a related note, linear filtering used widely to minimize the effect of noise have been shown to introduce marked distortion of the phase-space geometry of time series from chaotic systems[13]. Takens embedding procedure[14,15] provided an elegant way to reconstruct the multi-dimensional phase-space representation of nonlinear dynamical systems from their univariate time series representation using an appropriate

Center for Oral and Systemic Health, Marshfield Clinic Research Institute, 1000 North Oak Avenue, Marshfield, WI, 54449, USA. Correspondence and requests for materials should be addressed to R.N. (email: nagarajanr@marshfieldresearch.org)

time delay and embedding dimension[15,16]. It was perhaps one of the primary drivers in investigating the presence of deterministic chaos from time series realizations. Subsequently, an array of approaches with ability to provide insight into the generative mechanism behind a given time series under the broad theme "surrogate testing" were proposed. Surrogate testing is similar to statistical resampling techniques[17] and used widely to investigate the presence of dynamical nonlinearities in experimental time series[18–25]. On a related note, dynamical nonlinearities are an essential ingredient of deterministic chaotic processes. There have been several noteworthy contributions to surrogate testing from the statistical physics community[26–33] summarized in recent reviews[27,34].

Essential ingredients of classical surrogate testing include (a) an empirical time series sample, (b) null hypothesis, (c) discriminant statistic or dynamical invariant, (d) surrogate generation algorithm and (e) a statistical test. The empirical sample has traditionally been a single time series realization from the given system of interest. The null hypothesis assumes the generative mechanism of the given empirical sample. Surrogate algorithms are designed to generate time series realizations (i.e. surrogates) from the given empirical sample retaining critical properties that align with the null hypothesis. For these reasons, surrogates are also regarded as constrained randomized realizations[27,35]. Several surrogate generation algorithms have been proposed in literature. These include (a) Random Shuffled Surrogates, (b) Phase-Randomized Surrogates (Fourier Transform Surrogates, FT)[26], (c) Amplitude Adjusted Fourier Transform Surrogates (AAFT) and (d) Iterated Amplitude Adjusted Fourier Transform (IAAFT)[26–28]. Each of these surrogate algorithms addresses a particular null hypothesis. Random shuffled surrogate investigates whether the given empirical sample is uncorrelated noise and retains the probability distribution of the empirical sample in the surrogate realization destroying the correlation in the empirical sample. Thus any discriminant statistic sensitive to the correlation in the given data can be used as a discriminant statistic. FT surrogates preserve the power-spectrum of the given empirical sample in the surrogate realizations by constrained randomization of the phases. As noted earlier, preserving the power-spectrum is sufficient to determine the optimal parameters of linearly correlated processes. FT surrogates can be used to investigate the presence of nonlinear correlation in the given empirical sample but does not provide insight into the nature of nonlinearity. Thus any discriminant statistic sensitive to nonlinear correlations is a reasonable choice for FT surrogates. Subsequently, AAFT surrogates[26] were proposed in order to address the null hypothesis that the given process is a static, invertible nonlinear transform of a linearly correlated noise by following a phase-randomization and rank ordering procedure. IAAFT surrogates[28] has been shown to preserve the spectrum as well as the probability distribution of the given empirical sample in the surrogate realization while overcoming the flatness bias prevalent in AAFT surrogates. The primary objective of IAAFT surrogates was to identify potential dynamical nonlinearities in the given time series. Thus any discriminant statistic sensitive to dynamical nonlinearities (e.g. dynamical invariants) can be used for AAFT and IAAFT surrogates. Several additional surrogate algorithms have also been proposed since then[34]. However, surrogates in the present study are generated using the IAAFT surrogates. Finally, parametric and non-parametric statistical tests were proposed to assess significant difference in the discriminant statistic estimates between the empirical sample and the surrogate counterparts[27].

Traditional surrogate testing approaches while helpful have inherent limitations. They primarily rely on statistical comparison of discriminant statistic estimates on a single representative sample (i.e. empirical sample) to those obtained on their corresponding surrogate realizations, Fig. 1a. While the choice of empirical sample can be attributed to implicit ergodic assumptions[36], generating long time series so as to enable robust estimation of dynamical invariants and discriminant statistics can be especially challenging in experimental settings as it demands controlling a number of factors. Experimental time series such as those from physiological systems have been especially known to exhibit variations between subjects within a given disease group or cohort. These in turn encourages accommodating multiple realizations as opposed to a single empirical sample in the surrogate testing framework for enhanced generalizability of the findings. In such a scenario, each realization can be paired with the corresponding surrogate realization, Fig. 1b. As in the case of single empirical sample, if the multiple time series realizations are sufficiently long then it might be possible to statistically compare the distribution of discriminant statistic estimates on the given cohort to those estimated on its paired surrogate realizations addressing the null hypothesis that there is no significant difference in the discriminant estimates between the cohort and its surrogate counterpart, Fig. 1b. The present study takes a different tack to the classical surrogate testing. Its significance can be attributed to the following reasons. (a) The present study proposes a binary classification framework that uses a simple recurrent neural network with the raw time series as the input obviating the need to choose or estimate discriminant statistics or dynamical invariants. This is especially helpful across small lengths such as those discussed in the present study (L = 32, 64, 128) where estimation of discriminant statistics[37] can be challenging and unreliable. (b) It poses the classical statistical surrogate testing Fig. 1a,b, as a binary classification problem, Fig. 1c, using recurrent neural networks (RNN), Fig. 2, where the two classes of interest correspond to the multiple time series realizations from a given cohort and their corresponding IAAFT surrogate counterparts. Generalizability of the proposed approach is established by demonstrating the classifier performance on an independent validation data. (c) The results are demonstrated on short time series of lengths (L = 32, 64, 128) generated by nonlinear deterministic processes in chaotic regimes, nonlinear transforms of linearly correlated noise with varying parameters as well as experimental time series data.

Results

Accuracy of the binary classification framework was investigated across nonlinear deterministic, experimental time series and nonlinear transform of linearly correlated noise (Sec. Methods) with lengths (L = 32, 64, 128), Fig. 3. Only length (L = 128) was considered for the epileptic seizure in order to faithfully represent at least a few cycles of the seizure dynamics. Convergence of RNN training and validation loss for representative time series realizations is shown in Fig. 4. Accuracy of the test data as a function of the epochs for each of these time series are shown in Figs 5–7 respectively. Representative accuracies for each of these data sets chosen from the plateau region of the plots where the training and validation loss were consistently low are enclosed in Table 1.
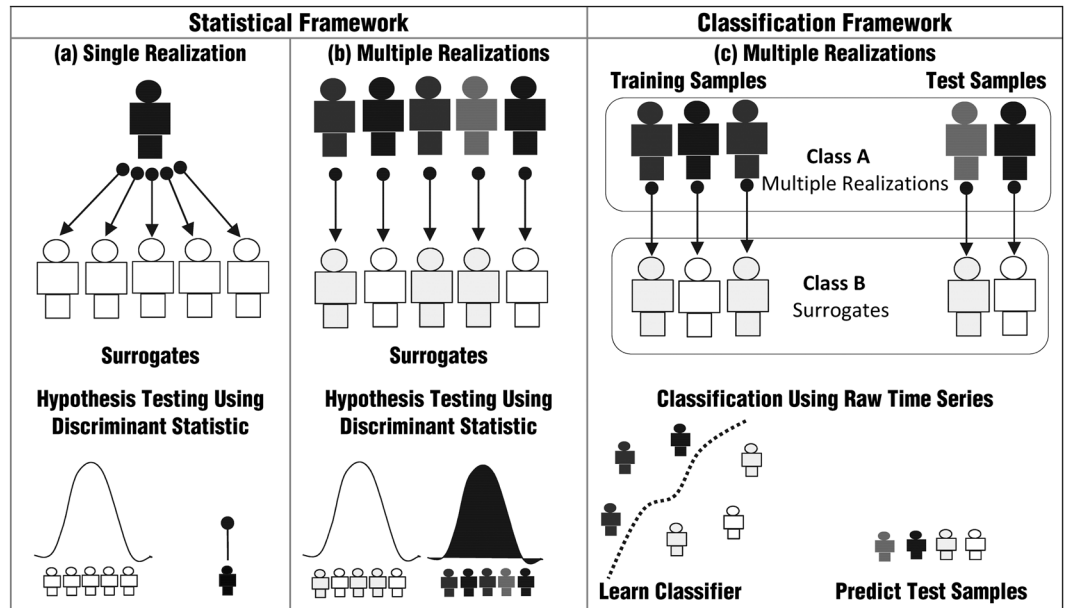
Figure 1. Classical surrogate testing approach on comparing the discriminant statistic estimate of a single empirical sample (black) to those obtained on their IAAFT surrogates (white) is shown in (**a**). Statistical hypothesis testing on comparing the discriminant statistic estimates of multiple realizations (shades of black) representing potential variation between them to those of its paired IAAFT surrogates (shades of white) is shown in (**b**). A binary classification framework using the raw time series from multiple realizations (shades of black) and those of its IAAFT surrogate counterpart (shades of white) is shown in (**c**). The dotted line in (**c**) represents the decision boundary of the classifier separating the multiple realizations (Class A) and its IAAFT surrogate counterparts (Class B). Performance and generalizability is subsequently demonstrated by the predictive ability of the classifier on an independent set of test samples.
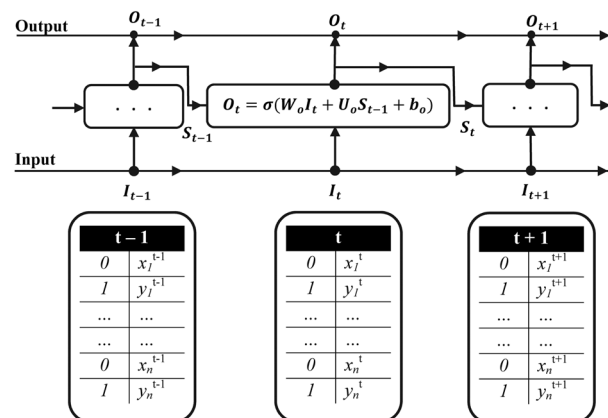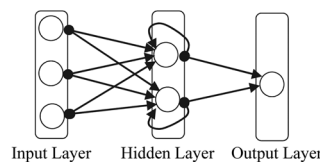

Figure 2. RNN cell unrolled in time representing the mapping between the input I_t, state of the network S_t and the output O_t through W_o, U_o and b_o is shown in (**a**). The activation function and the bias term are represented by σ and b_o respectively. The labels of the empirical sample x^t and the IAAFT surrogate realizations y^t at time t are represented by 0 and 1 respectively. A typical RNN architecture comprising of multiple RNN Cells in the hidden layer along with the input and output layers is shown in (**b**).
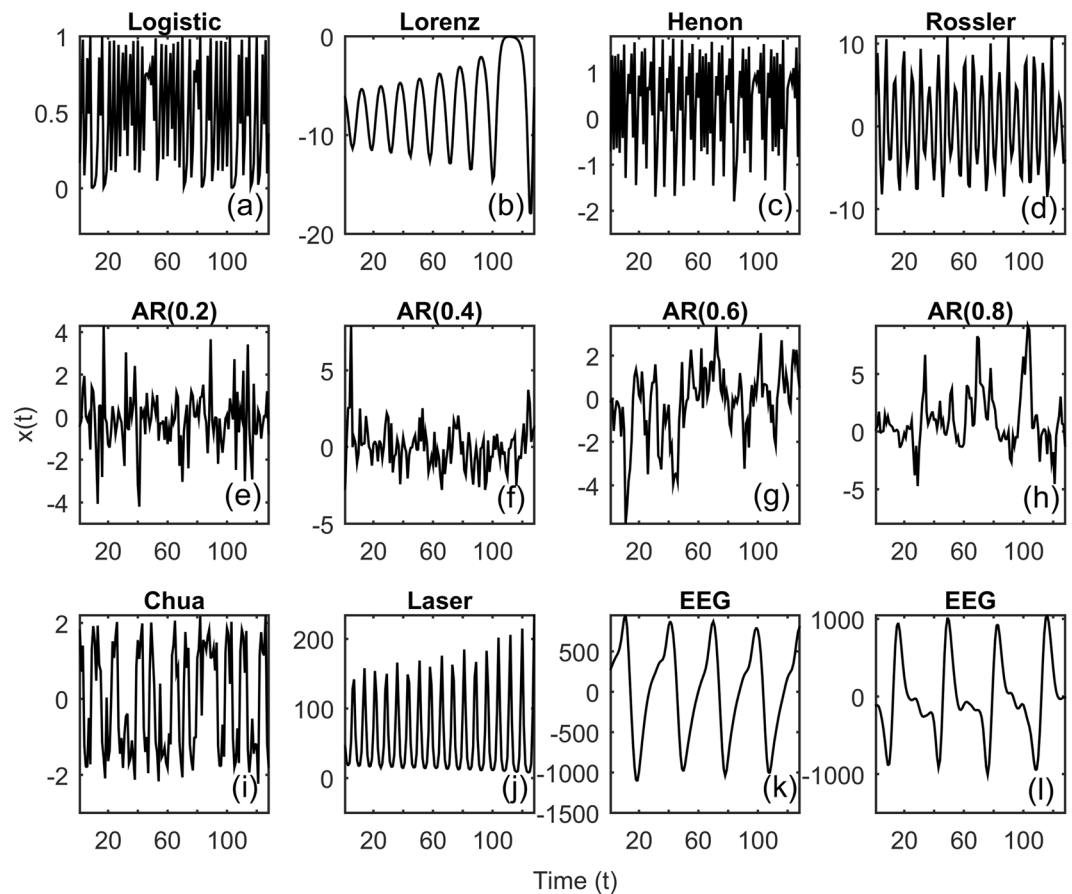
Figure 3. Representative time series realizations $x(t)$ of length (L = 128) as a function of time t for nonlinear deterministic processes (top), nonlinear transform of correlated noise (middle) and experimental time series (bottom) is shown. The top row represents nonlinear deterministic processes in chaotic regimes (Logistic Map, Lorenz Attractor, Henon Map, Rossler Attractor). Middle row represents nonlinear transform of linearly correlated noise with parameters ($\alpha = 0.2$, 0.4, 0.6 and 0.8). The bottom row represents time series realizations from experimental systems in chaotic regimes (Chua's Circuits, Santa Fe Laser) and EEG data during epileptic seizure.

Nonlinear deterministic process. For time series generated from discrete and continuous nonlinear deterministic systems (Logistic, Henon, Lorenz and Rossler, Sec. Methods), the accuracy of the classifier showed a marked transition towards larger values from 0.5 as a function of the epochs, Fig. 5. A one-sample binomial test rejected the null that the accuracy was similar to that of random guess (0.5) at a significance level ($\alpha = 0.05$), Table 1. These results were consistently observed across the three sample sizes (L = 32, 64, 128) and across the data sets demonstrating the classifiers ability to discern dynamical nonlinearities and their IAAFT surrogate counterparts. The number of neurons in the hidden layer of the RNN was fixed at (N = 10). The RNN parameters (Sec. Methods) were fixed across these data sets, Table 1.

Experimental time series. Experimental time series generated using Chua's circuits (L = 32, 64, 128) and Santa Fe Laser Time Series (L = 32, 64, 128) in chaotic regimes (Sec. Methods) exhibited accuracies much greater than 0.5, Fig. 6, as observed in the case of the nonlinear deterministic processes, Fig. 5. A one-sample binomial test rejected the null hypothesis that the representative accuracy was similar to that of random guess (0.5) at a significance level ($\alpha = 0.05$), Table 1. For the time series generated from Chua's circuits and the Santa Fe laser time series, the number of neurons in the hidden layer of the RNN were chosen as 20 for (L = 32, 64) and 25 for (L = 128), Table 1. All other parameters of the RNN were retained as discussed in (Sec. Methods). Three representative EEG signals of lengths (L = 128) during seizure from a recent study[3] were reinvestigated using the proposed approach. Unlike Chua's circuits and Santa Fe time series, it is important to note that the underlying process generating the EEG signals during seizures is unknown. However, several studies have investigated nonlinear dynamical aspects of seizures and the evolution of characteristic synchronization patterns accompanying seizures[38,39]. The accuracy of the classifier as a function of the epoch exhibited a marked transition from 0.5 for the EEG. A one-sample binomial test rejected the null that the representative accuracy was similar to that of random guess (0.5) at a significance level ($\alpha = 0.05$), Table 1. The number of neurons in the hidden layer of the RNN was fixed at (N = 20) for the three EEG signals, Fig. 6. All other parameters of the RNN were retained as discussed in (Sec. Methods).
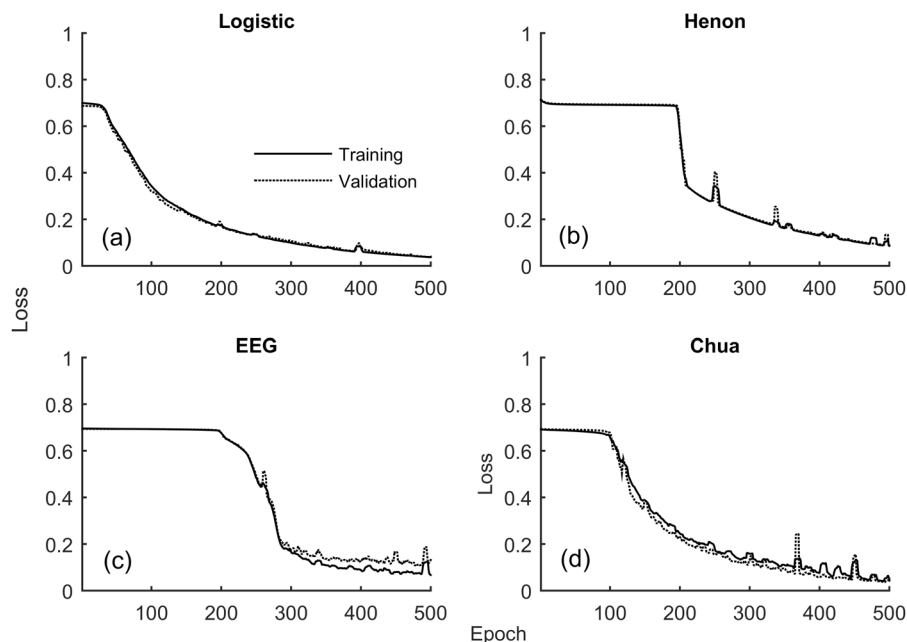
Figure 4. Training (solid line) and validation loss (dotted lines) of four representative time series (Logistic Map, Henon Map, EEG and Chua's Circuits) with length ($L = 128$) as a function of the epochs is shown.
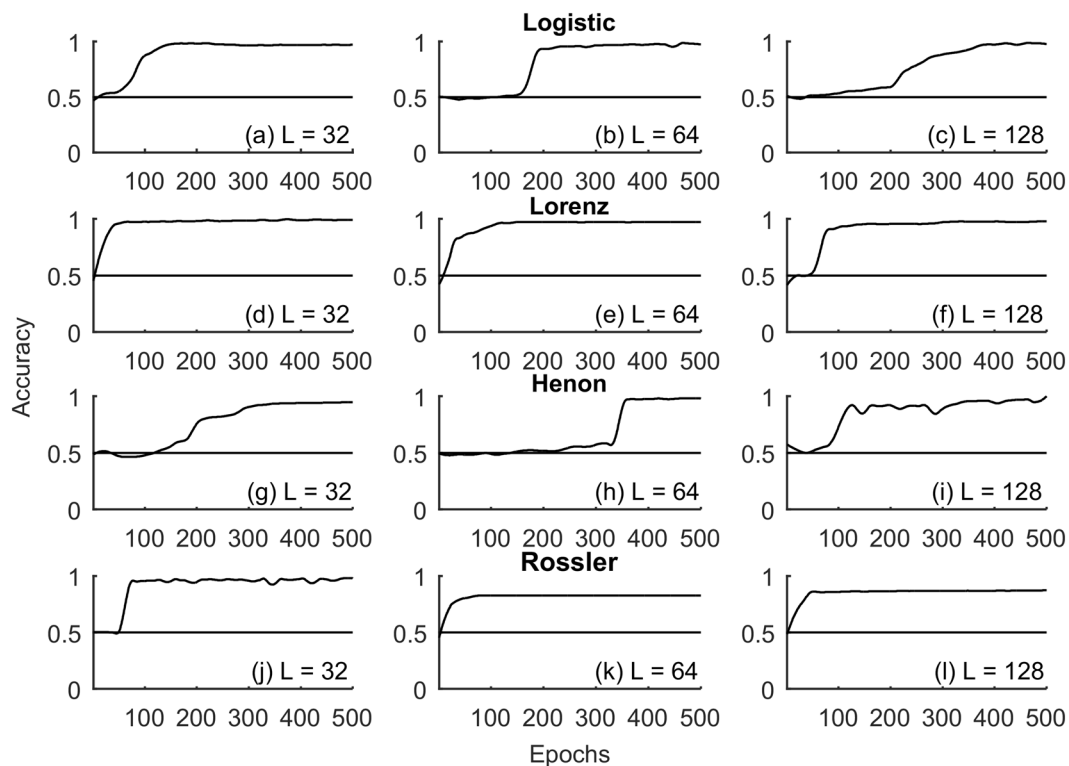

Figure 5. Accuracy of the classifier for the chaotic time series data generated for discrete and continuous dynamical systems (Logistic Map, Lorenz System, Henon Map, Rossler System) as a function of the number of epochs for lengths ($L = 32$, 64 and 128) is shown in each of the rows. The horizontal solid line corresponds to accuracy 50% characteristic of random guess and shown as a reference in each of the subplots.

Nonlinear transform of linearly correlated noise. Time series generated from a static nonlinear transform of linearly correlated noise[28] (Sec. Methods) were investigated with varying process parameters ($\alpha = 0.2$, 0.4, 0.6, 0.8) in the stationary regime, Fig. 7. Unlike the case of nonlinear deterministic chaos, accuracy estimates from
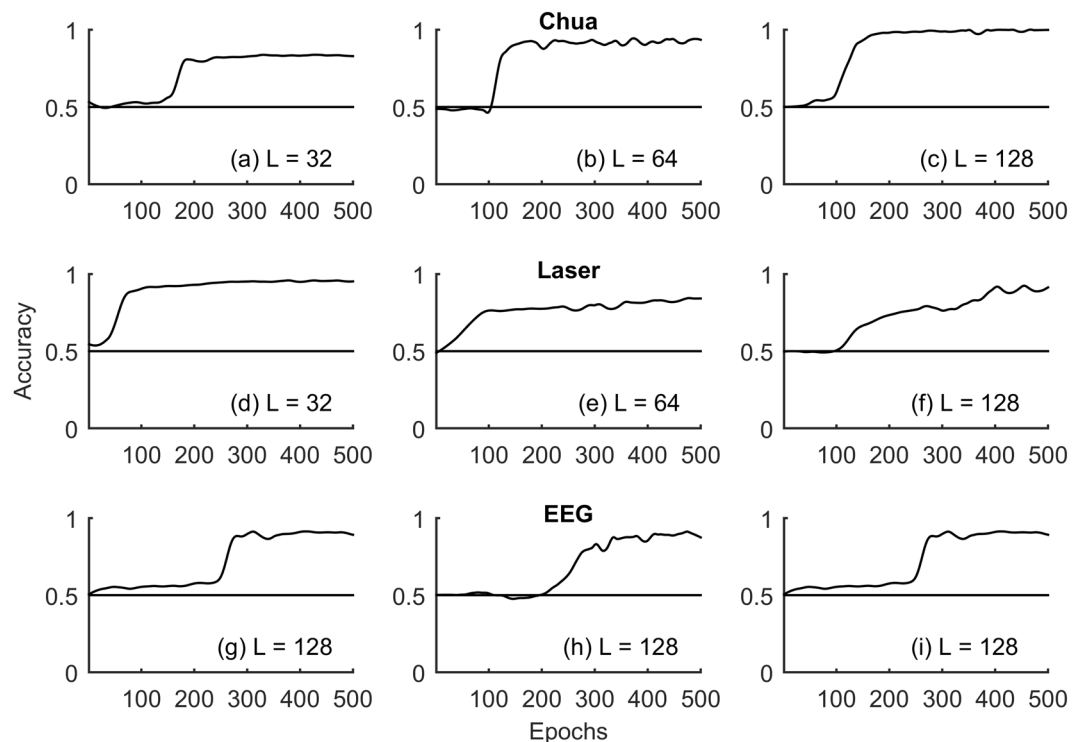
Figure 6. Accuracy of the classifier for the experimental time series data generated from Chua's circuits in chaotic regime (L = 32, 64, 128), Santa Fe Laser time series in chaotic regime (L = 32, 64, 128) and three representative EEG signals during seizure (L = 128) as a function of the epochs is shown. The number of neurons in the hidden layer for each of these cases is enclosed in Table 1. The horizontal solid line corresponds to accuracy 50% characteristic of random guess and shown as a reference in each of the subplots.

Time Series	N = 32	N = 64	N = 128
Nonlinear Deterministic Processes			
Logistic (10 neurons)	0.98*	0.97*	0.96*
Henon (10 neurons)	0.93*	0.96*	0.87*
Lorenz (10 neurons)	0.98*	0.98*	0.97*
Rossler (10 neurons)	0.94*	0.82*	0.83*
Experimental Time Series Data			
Santa Fe Laser Time Series	0.94* (20 neurons)	0.84* (20 neurons)	0.91* (25 neurons)
Chua's Oscillator	0.82* (20 neurons)	0.92* (20 neurons)	0.98* (25 neurons)
Epileptic Seizure 1		0.95* (20 neurons)	
Epileptic Seizure 2		0.95* (20 neurons)	
Epileptic Seizure 3		0.98* (20 neurons)	
Nonlinearly Correlated Noise			
($\alpha = 0.2$) (10 neurons)	0.53	0.48	0.52
($\alpha = 0.4$) (10 neurons)	0.49	0.50	0.51
($\alpha = 0.6$) (10 neurons)	0.51	0.46	0.49
($\alpha = 0.8$) (10 neurons)	0.49	0.53	0.50

Table 1. Classification accuracies for nonlinear deterministic processes in chaotic regimes, experimental time series and non-deterministic processes. Accuracy estimates that were statistically significant ($\alpha = 0.05$) from 0.5 in a one-sample binomial test are shown by asterisk.

the RNN classification framework did not show an appreciable change from that of random guess (0.5), Fig. 7 as expected, indicating that the properties of the given data are not significantly different from those of their IAAFT surrogate counterparts. A one-sample binomial test did not reject the null that the representative accuracy was similar to that of random guess (0.5) at a significance level ($\alpha = 0.05$), Table 1. These results were consistent across the different process parameters ($\alpha = 0.2, 0.4, 0.6, 0.8$) and lengths (L = 32, 64, 128). The number of neurons in the hidden layer of the RNN were fixed at (N = 10) similar to that of the nonlinear deterministic processes. Any further increase in the number of neurons in the hidden layer resulted in overfitting like behavior accompanied
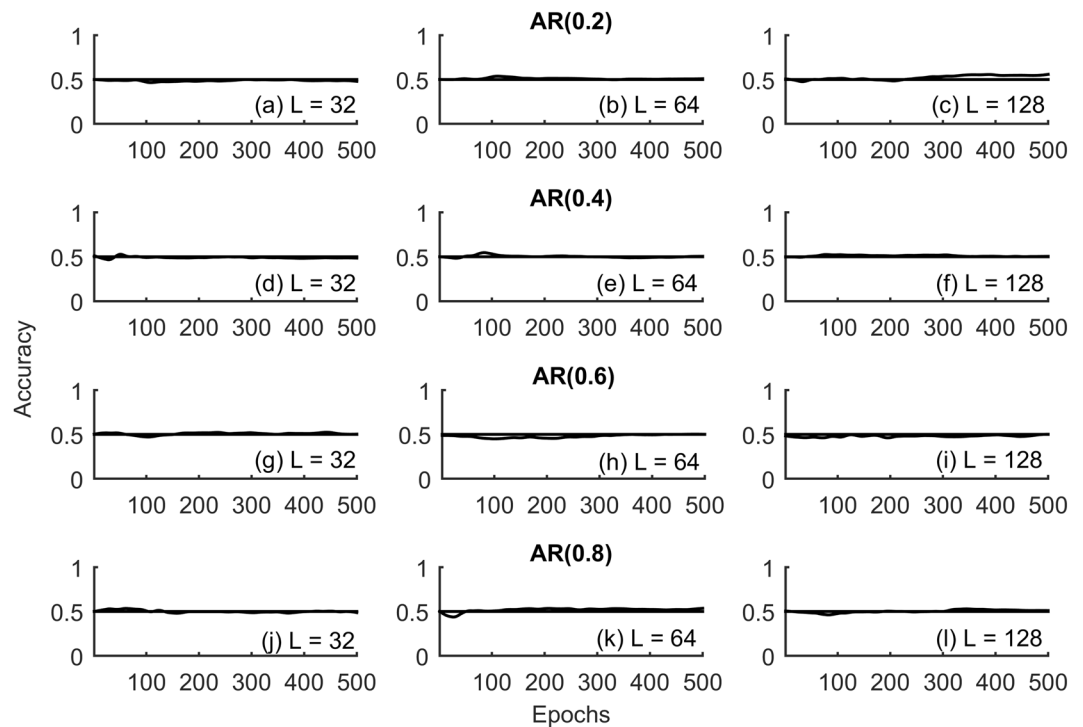
Figure 7. Accuracy of the classifier for the time series data generated from static nonlinear transform of linearly correlated noise with parameters ($\alpha = 0.2, 0.4, 0.6, 0.8$) as a function of the number of epochs for lengths ($L = 32, 64$ and 128) is shown in each of the rows. The number of neurons in the hidden layer for each of these cases ($N = 10$) is enclosed in Table 1. The horizontal solid line corresponds to accuracy 50% characteristic of random guess and shown as a reference in each of the subplots.

by marked separation in the training and validation loss. All other parameters of the RNN were retained as discussed in (Sec. Methods).

Discussion

Several studies have successfully used surrogate testing techniques to discern static and dynamical nonlinearities such as those from deterministic chaotic systems. Their ability to provide insights into the generative mechanism from the given time series realization(s) is a primary reason for their widespread adoption across a spectrum of disciplines. Traditional surrogate testing while helpful has inherent limitations. It subscribes to statistical hypothesis testing and investigates the separation of a chosen discriminant statistic or dynamical invariant between the given empirical sample and its surrogate counterpart. These discriminant statistic and dynamical invariants essentially capture certain facets of the given time series and their choice can be non-trivial with marked impact on the conclusions. Dynamical invariants and discriminant statistic estimation can be especially challenging across short time series such as those discussed in the present study. The proposed approach obviates the need to estimate discriminant statistics or dynamical invariants and uses the raw time series in the surrogate testing procedure. Conclusions based on traditional surrogate testing are also based on single realization or empirical sample. However, drawing conclusions based on a single realization can be a limitation from a practical standpoint. This is especially true with experimental data such as those from physiological systems and healthcare settings where variations are common within a given cohort. These in turn demand incorporation of multiple realizations for enhanced generalizability with potential to assist in clinical decision making. The proposed approach accommodates multiple realizations simultaneously and poses the traditional statistical hypothesis testing framework as a classification framework. For the nonlinear deterministic process, a marked increase in accuracy was observed as a function of epochs unlike that of the non-deterministic processes. Ideally, the error rate (i.e. 1 – accuracy) distribution may be positively skewed for large number of epochs for the nonlinear deterministic whereas that of non-deterministic process is expected to be relatively uniform.

Generating long stationary time series from experimental systems can be challenging as it demands controlling a number of factors for extended periods. The present study provides a suitable alternative by using multiple short time series realizations, hence expected to find wide applications across a number of settings. While the results presented in this study investigated the performance of a simple RNN with 10–20 neurons and a single hidden layer, the RNN hyperparameters in general will have to be tuned. The results presented showed a marked increase in accuracy across the dynamical nonlinearities generated from nonlinear deterministic processes in chaotic regimes. However, it is important to note that dynamical nonlinearities can arise across deterministic as well as non-deterministic settings. The latter would include deterministic dynamical systems with dynamical and measurement noise. Therefore, conclusions on the presence of dynamical nonlinearities do not necessarily imply presence of deterministic chaos.

Methods

Working principle of the IAAFT Algorithm. The IAAFT algorithm[28] is an iterative procedure that aims to retain the power-spectrum as well as the distribution of the given empirical sample in the surrogate realizations. As noted earlier, retaining the power-spectrum retains the linear characteristics of the time series. Rank ordering aspect of IAAFT is useful in retaining static, invertible nonlinearities but not the dynamical nonlinearities in the given empirical sample. The working principle of IAAFT is enclosed below for completeness, a detailed explanation and implementation can be found in the following references[24,27,28,34,40].

Let the given empirical sample be $\{x_n\}$.

> **Step 1:** Generate a random shuffle $\{x_n^i\}$ of the given empirical sample $\{x_n\}$.
> **Step 2:** Preserving the power spectrum in the surrogate.
> Generate the Fourier transform of $\{x_n\}$ and $\{x_n^i\}$. Let the corresponding squared amplitudes be $\{S_k^2\}$ and $\{S_k^{2i}\}$ respectively. Substitute $\{S_k^{2i}\}$ by $\{S_k^2\}$ and generate the inverse Fourier transform to obtain $\{y_n\}$.
> **Step 3:** Preserving the distribution in the surrogate.
> Rank order $\{y_n\}$ to have same distribution as $\{x_n\}$ resulting in the surrogate $\{x_n^{i+1}\}$.
> **Step 4:** Repeat Steps 2 and 3 so as to minimize the discrepancy in the spectrum between empirical sample and its surrogate.

Nonlinear deterministic process. Time series were generated from discrete and continuous dynamical systems in chaotic regimes. Representative time series in chaotic regimes is shown in Fig. 5. Time series data for the continuous dynamical systems were generated using explicit Runge-Kutta (4, 5) implemented as a part of the MATLAB ode45 function[41].

(i) *Logistic map in chaotic regime* $(r = 4.0)$[42],

$$x_{t+1} = rx_t(1 - x_t)$$

(ii) *Henon map in chaotic regime* $(\alpha = 1.4, \beta = 0.3)$[43,44],

$$x_{t+1} = 1 - \alpha x_t^2 + y_t$$
$$y_{t+1} = \beta x_t$$

(iii) *Lorenz system in chaotic regime* $(\sigma = 10, \ \rho = 28, \ \beta = 8/3)$[45]

$$\frac{dx}{dt} = \sigma(y - x)$$
$$\frac{dy}{dt} = x(\rho - z) - y$$
$$\frac{dz}{dt} = xy - \beta z$$

(iv) *Rossler system in chaotic regime* $(\alpha = 0.2, \beta = 0.2, \gamma = 5.7)$[46],

$$\frac{dx}{dt} = -y - z$$
$$\frac{dy}{dt} = x + \alpha y$$
$$\frac{dz}{dt} = \beta + z(x - \gamma).$$

Experimental time series data.

(i) *Chua's Circuit*
Chua's circuit[2,47] is a simple autonomous electric circuit and can be readily designed using resistors, capacitors, inductors and a nonlinear element. It is perhaps one of the most popular experimental evidence of deterministic chaos. An equivalent dimensionless model with parameters $(\alpha = 15.6, \ \beta = 28, \ m_0 = -8/7, \ m_1 = -5/7)$ has also been proposed in literature to capture the behavior of the original circuit[2,47].

$$\frac{dx}{dt} = \alpha(y - x - f(x))$$
$$\frac{dy}{dt} = x - y + z$$
$$\frac{dz}{dt} = -\beta y$$

where the piece-wise linear function $f(x) = m_1 x + 0.5(m_0 - m_1)(|x + 1| - |x - 1|)$.

(ii) *Santa-Fe Laser Time Series*

Several studies have provided compelling evidence of chaos across distinct laser systems[48–50]. The present study re-investigates Santa Fe Laser time series of 1000 samples derived from a Far-Infrared (FIR) laser in chaotic regime[51,52].

(iii) *Epileptic Seizure Time Series*

Electroencephalograms (EEG) signals recorded during epileptic seizure have been argued to exhibit patterns characteristic of nonlinear dynamical processes. Three representative EEG samples from seizure subjects reported in a recent study[3] were re-investigated using the proposed classification framework. As recommended in the original study[3], the three EEG signals were pre-processed using a 4th order low-pass Butterworth filter[1] to minimize the impact of noise and impose the high-frequency cut-off at 40 Hz. In order to capture a few cycles of the EEG waveform only samples with length (N = 128) were investigated.

Nonlinear transform of linearly correlated noise.

$$x_t = \alpha x_{t-1} + \epsilon_t; \ y_t = x_t \sqrt{|x_t|};$$

The above example was motivated by a recent study[28]. The process x_t is a linearly correlated noise where ϵ_t is zero-mean, unit variance normally distributed uncorrelated noise with y_t representing a static nonlinear transform of x_t. Several choices of the process parameters ($\alpha = 0.2, 0.4, 0.6, 0.8$) were investigated in the present study. Representative time series data generated by nonlinear transform of linearly correlated noise with process parameters ($\alpha = 0.2, 0.4, 0.6, 0.8$) is shown in Fig. 3.

Surrogate testing using a recurrent neural network. *Data.* The time series realizations was fixed at (N = 1000) across all the data sets. Time series of three different lengths (L = 32, 64, 128) were investigated. For the experimental data sets in the present study, (N = 1000) realizations was generated by randomly choosing a sequence of time points of length (L = 32, 64, 128) from the given data. Representative samples of the various time series are shown in Fig. 3.

RNN. RNN architectures by very design are ideal for prediction and classification of sequence data. RNN cell unfolded in time[53,54] and a typical RNN architecture comprising of multiple RNN cells in the hidden layer is shown in Fig. 2. In the present study, the input and output of the RNN were the time series realizations and their corresponding labels respectively. The time series realizations (N = 1000) was split into training samples (75%) and test samples (25%). Since each time series realization was paired to its IAAFT surrogate counterpart, the classes were balanced by very design justifying the choice of accuracy as a classifier performance measure in the present study. RNN parameters were chosen after experimentation[55]. RNN was implemented using Keras high-level neural network API with Tensorflow backend[53,54] and Adam optimizer (ADAM)[56] (learning rate 0.0001, batch size 16 and binary cross-entropy loss) for the data sets in the present study. The number of neurons for the synthetic data sets generated from nonlinear dynamical systems, was chosen as (N = 10), Table 1. For the nonlinearly correlated noise, the number of neurons was also fixed at (N = 10), Table 1. For the experimental time series data, the number of hidden neurons varied and enclosed in Table 1. Neurons in the hidden layer were accompanied by rectified linear unit (ReLU) activation function whereas those in the output layer had sigmoid activation function. RNN learning curves were inspected during the training phase for potential overfitting. The validation split in the training phase was set at 30%, implying the last 30% of the training data were used as internal validation in computing the accuracy and loss curves as a function of the epoch. The training and validation loss as a function of the epoch for representative nonlinear deterministic processes and experimental time series are shown in Fig. 4. As can be observed for each of these cases, the training and validation loss simultaneously transitioned to markedly lower values with increasing epochs. While certain RNN applications do encourage having a validation loss lower than that of the training loss, the present study estimated the accuracies (Table 1) at the epoch where the training and validation loss were simultaneously low, Fig. 4. A smoothing window of five samples was used to generate the learning curves, Fig. 4, and accuracy profiles, Figs 5–7, as a function of the epochs.

Data Availability

The experimental data sets used in present study are publicly available and the corresponding references are provided. The equations to the synthetic data sets are provided as a part of the manuscript. All implementations and figures were done in MATLAB. RNN implementation was accomplished using the open-source package Keras. The surrogate generation algorithms have been implemented as a part of the (TISEAN: TIme SEries ANalysis) package MATLAB package (MATS: Measures of Analysis of Time Series). The references to these packages and the experimental data are included in the manuscript.

References

1. Proakis, J. G. & Manolakis, D. G. *Digital signal processing* (*3rd ed.*): *principles, algorithms, and applications*. (Prentice-Hall, Inc. 1996).
2. Chua, L., Komuro, M. & Matsumoto, T. The double scroll family. *IEEE transactions on circuits and systems* **33**, 1072–1118 (1986).
3. Andrzejak, R. G. *et al.* Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E* **64**, 061907 (2001).
4. Zhang, D., Györgyi, L. & Peltier, W. R. Deterministic chaos in the Belousov–Zhabotinsky reaction: Experiments and simulations. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **3**, 723–745 (1993).
5. Ghosh, S. *et al.* Experimental evidence of intermittent chaos in a glow discharge plasma without external forcing and its numerical modelling. *Physics of Plasmas* **21**, 032303 (2014).
6. Matsumoto, T. Chaos in electronic circuits. *Proceedings of the IEEE* **75**, 1033–1057 (1987).
7. Kauffman, S. A. *The origins of order: Self-organization and selection in evolution*. (OUP USA, 1993).
8. Strogatz, S. H. *Nonlinear Dynamics and Chaos with Student Solutions Manual: With Applications to Physics, Biology, Chemistry, and Engineering*. (CRC Press, 2018).
9. Ott, E., Grebogi, C. & Yorke, J. A. Controlling chaos. *Physical Review Letters* **64**, 1196–1199, https://doi.org/10.1103/PhysRevLett.64.1196 (1990).
10. Ditto, W. L., Rauseo, S. N. & Spano, M. L. Experimental control of chaos. *Physical Review Letters* **65**, 3211–3214, https://doi.org/10.1103/PhysRevLett.65.3211 (1990).
11. Ditto, W. & Munakata, T. Principles and applications of chaotic systems. *Communications of the ACM* **38**, 96–102 (1995).
12. Farmer, D., Crutchfield, J., Froehling, H., Packard, N. & Shaw, R. Power spectra and mixing properties of strange attractors. *Annals of the New York Academy of Sciences* **357**, 453–471 (1980).
13. Theiler, J. & Eubank, S. Don't bleach chaotic data. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **3**, 771–782 (1993).
14. Takens, F. In *Dynamical systems and turbulence, Warwick 1980* 366–381 (Springer, 1981).
15. Sauer, T., Yorke, J. A. & Casdagli, M. Embedology. *Journal of statistical Physics* **65**, 579–616 (1991).
16. Kennel, M. B., Brown, R. & Abarbanel, H. D. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical review A* **45**, 3403 (1992).
17. Efron, B. *The jackknife, the bootstrap, and other resampling plans*. Vol. 38 (Siam, 1982).
18. Nagarajan, R., Szczepanski, J. & Wajnryb, E. Interpreting non-random signatures in biomedical signals with Lempel–Ziv complexity. *Physica D: Nonlinear Phenomena* **237**, 359–364 (2008).
19. Nagarajan, R. Surrogate testing of linear feedback processes with non-Gaussian innovations. *Physica A: Statistical Mechanics and its Applications* **366**, 530–538 (2006).
20. Govindan, R., Narayanan, K. & Gopinathan, M. On the evidence of deterministic chaos in ECG: Surrogate and predictability analysis. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **8**, 495–502 (1998).
21. Shiogai, Y., Stefanovska, A. & McClintock, P. V. E. Nonlinear dynamics of cardiovascular ageing. *Physics reports* **488**, 51–110 (2010).
22. Paluš, M. & Stefanovska, A. Direction of coupling from phases of interacting oscillators: an information-theoretic approach. *Physical Review E* **67**, 055201 (2003).
23. Kugiumtzis, D. & Larsson, P. In *Chaos in Brain?* 329–332 (World Scientific, 2000).
24. Kugiumtzis, D. & Tsimpiris, A. Measures of analysis of time series (MATS): a MATLAB toolkit for computation of multiple measures on time series data bases. *arXiv preprint arXiv:1002.1940* (2010).
25. Rapp, P., Cellucci, C. J., Watanabe, T., Albano, A. & Schmah, T. Surrogate data pathologies and the false-positive rejection of the null hypothesis. *International Journal of Bifurcation and Chaos* **11**, 983–997 (2001).
26. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Farmer, J. D. Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena* **58**, 77–94 (1992).
27. Schreiber, T. & Schmitz, A. Surrogate time series. *Physica D: Nonlinear Phenomena* **142**, 346–382 (2000).
28. Schreiber, T. & Schmitz, A. Improved surrogate data for nonlinearity tests. *Physical Review Letters* **77**, 635 (1996).
29. Kugiumtzis, D. Surrogate data test for nonlinearity including nonmonotonic transforms. *Physical Review E* **62**, R25 (2000).
30. Paluš, M. Testing for nonlinearity using redundancies: Quantitative and qualitative aspects. *Physica D: Nonlinear Phenomena* **80**, 186–205 (1995).
31. Kugiumtzis, D. Test your surrogate data before you test for nonlinearity. *Physical Review E* **60**, 2808 (1999).
32. Prichard, D. & Theiler, J. Generating surrogate data for time series with several simultaneously measured variables. *Physical review letters* **73**, 951 (1994).
33. Rapp, P., Albano, A., Zimmerman, I. & Jimenez-Montano, M. Phase-randomized surrogates can produce spurious identifications of non-random structure. *Physics letters A* **192**, 27–33 (1994).
34. Lancaster, G., Iatsenko, D., Pidde, A., Ticcinelli, V. & Stefanovska, A. Surrogate data for hypothesis testing of physical systems. *Physics Reports* (2018).
35. Theiler, J. & Prichard, D. Constrained-realization Monte-Carlo method for hypothesis testing. *Physica D: Nonlinear Phenomena* **94**, 221–235 (1996).
36. Eckmann, J.-P. & Rueiie, D. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics* **57** (1985).
37. Rapp, P. E., Albano, A. M., Schmah, T. & Farwell, L. Filtered noise can mimic low-dimensional chaotic attractors. *Physical review E* **47**, 2289 (1993).
38. Lehnertz, K. Epilepsy and nonlinear dynamics. *Journal of biological physics* **34**, 253–266 (2008).
39. Jiruska, P. *et al.* Synchronization and desynchronization in epilepsy: controversies and hypotheses. *The Journal of physiology* **591**, 787–797 (2013).
40. Hegger, R., Kantz, H. & Schreiber, T. Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **9**, 413–435 (1999).
41. Shampine, L. F. & Reichelt, M. W. The matlab ode suite. *SIAM journal on scientific computing* **18**, 1–22 (1997).
42. May, R. M. Simple mathematical models with very complicated dynamics. *Nature* **261**, 459 (1976).
43. Henon, M. A Two-dimensional Mapping with a Strange Attractor. *Commun. math. Phys* **50**, 69–77 (1976).
44. Hénon, M. Numerical study of quadratic area-preserving mappings. *Quarterly of applied mathematics*, 291–312 (1969).
45. Lorenz, E. N. Deterministic nonperiodic flow. *Journal of the atmospheric sciences* **20**, 130–141 (1963).
46. Rössler, O. E. An equation for continuous chaos. *Physics Letters A* **57**, 397–398 (1976).
47. Alligood, K. T., Sauer, T. D. & Yorke, J. A. *Chaos*. (Springer, 1996).
48. Weiss, C., Klische, W., Ering, P. & Cooper, M. Instabilities and chaos of a single mode NH3 ring laser. *Optics communications* **52**, 405–408 (1985).
49. Abraham, N. *et al.* In *Laser Physics* 107–131 (Springer, 1983).
50. Dupertuis, M.-A., Salomaa, R. & Siegrist, M. The conditions for Lorenz chaos in an optically-pumped far-infrared laser. *Optics communications* **57**, 410–414 (1986).
51. Weigend, A. S. *Time series prediction: forecasting the future and understanding the past*. (Routledge, 2018).
52. Weiss, C.-O., Hübner, U., Abraham, N. B. & Tang, D. Lorenz-like chaos in NH3-FIR lasers. *Infrared Physics & Technology* **36**, 489–512 (1995).
53. Chollet, F. Deep learning with Python (2018).

54. Chollet, F. & Allaire, J. Deep Learning with R (2018).
55. Bengio, Y. In *Neural networks: Tricks of the trade* 437–478 (Springer, 2012).
56. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

Additional Information