


Article

Criticality Analysis: Bio-Inspired Nonlinear Data Representation

Tjeerd V. olde Scheper 

School of Engineering, Computing and Mathematics, Oxford Brookes University, Wheatley Campus, Oxford OX33 1HX, UK; tvolde-scheper@brookes.ac.uk

Abstract: The representation of arbitrary data in a biological system is one of the most elusive elements of biological information processing. The often logarithmic nature of information in amplitude and frequency presented to biosystems prevents simple encapsulation of the information contained in the input. Criticality Analysis (CA) is a bio-inspired method of information representation within a controlled Self-Organised Critical system that allows scale-free representation. This is based on the concept of a reservoir of dynamic behaviour in which self-similar data will create dynamic nonlinear representations. This unique projection of data preserves the similarity of data within a multidimensional neighbourhood. The input can be reduced dimensionally to a projection output that retains the features of the overall data, yet has a much simpler dynamic response. The method depends only on the Rate Control of Chaos applied to the underlying controlled models, which allows the encoding of arbitrary data and promises optimal encoding of data given biologically relevant networks of oscillators. The CA method allows for a biologically relevant encoding mechanism of arbitrary input to biosystems, creating a suitable model for information processing in varying complexity of organisms and scale-free data representation for machine learning.

Keywords: bio-inspired computing; machine learning; self-organised criticality



Citation: olde Scheper, T.V. Criticality Analysis: Bio-Inspired Nonlinear Data Representation. *Entropy* **2023**, *25*, 1660. <https://doi.org/10.3390/e25121660>

Academic Editors: Anne Humeau-Heurtier and Javier Escudero

Received: 7 November 2023

Revised: 7 December 2023

Accepted: 12 December 2023

Published: 14 December 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

One of the aims of methodologies applied to feature extraction for machine learning is to find separation boundaries within the datasets such that categorisation algorithms can readily classify the samples. Different methods rely on specific approaches, based on probabilistic or deterministic principles. In general, the most effective approaches are those that emphasise the common features shared within the class, understate features that separate class members, and ideally perform the opposite for between-class comparisons [1]. Most commonly, these problems form part of some regression where multimodal input is reduced to a few representative variables [2]. Here, it is proposed that these types of problems are common in biological systems and that a bio-inspired approach based on simple connected networks of nonlinear controlled biochemical oscillators is capable of creating nonlinear representation spaces that facilitate classification greatly. The network is based on individual units from biochemical origins that are nonlinear and can exhibit chaotic behaviour. These units are then controlled using the Rate Control of Chaos (RCC) method devised by [3], which can stabilise the dynamic behaviour into steady state or periodic orbits. The method ensures that each individual unit can maintain stability even when perturbed externally. The cumulative effect of the total network shows stable dynamic behaviour that shares properties with Self-Organised Critical systems [4].

The concept of Self-Organised Criticality is the notion that a nonlinear dynamic system can be at an equilibrium or critical point such that a small perturbation may cause it to transition from one state to another state. The underlying dynamic behaviour of the elements that form such a system is based on local dynamic behaviour that contributes to global stable states. Such a mechanism has previously been conjectured to underlie

biological dynamic behaviour by [5,6]. This requires, however, a control method that is capable of ensuring the dynamic stability of those units; even if they are perturbed externally. These units would, without such a control method, result in unstable dynamic behaviours, such as is the case in spatiotemporal chaos. The use of RCC allows such a network to maintain regularity even when perturbed, and from this network, it has the emergent property that contains all three key features of a Self-Organised Critical system, namely non-trivial scaling due to external perturbation, spatiotemporal power-law correlation with respect to the total behaviour of the network, and self-tuning to a critical point [7]. The concept of stability and SOC despite external perturbations allows a representation of input resulting in a uniquely dynamic consistent representation which forms the basis of the Criticality Analysis method.

The resulting network of oscillators is adaptive, and dynamically stable, allowing a system that is therefore suited to act as a deterministic reservoir of oscillators, capable of receiving input in a deterministic and consistent manner that can be apparently scale-free. This reservoir of oscillators is, even in its most basic representation, capable of uniquely creating a stable and consistent representation of input, especially if this is in itself dynamic.

Two approaches that share similarity with the CA method are the Liquid State Machine (LSM) [8] and the Echo State Network (ESN) [9], which represent the use of a network of artificial neurons with random connections, perturbed by external data. These networks are then readout by an output layer of neurons that requires training using a suitable learning rule, which can be efficient [10]. Such a representation depends much on the randomised connectivity for a trainable response, may not produce uniquely represented datasets, can create spurious states, and, due to the decorrelation rules applied during training to ensure non-overlapping output, may lose information contained in the data [11]. The approach used for ESN may provide efficient representation at the edge of criticality [12]; however, this is still based on a Recurrent Neural Network approximation of the dynamic system to be represented. Furthermore, existing correlations are not guaranteed to be retained close to the edge and extensive tuning is required to drive the ESN to this edge.

The advantage of the Criticality Analysis method is that the network states are deterministic, in the sense that the resulting states are consistent with the presented input, and each input will result in the same state. Furthermore, similar input will result in nearby states in the dynamic sense, where the representation spaces are dynamically separated into spatial domains. The CA method will also provide the ability to change the dynamic behaviour of the network in frequency or other dynamic properties. Additionally, unlike in the case of an echo state machine or reservoir-computing approach [9], the number of units is fairly small, and the number of oscillators and their connectivity affects the dynamic response, but not so much the representation, allowing scale-free consistent representations, as will be shown in the results.

The proposed Criticality Analysis method can be recognised as a nonlinear mapping of the feature space into a lower dimensional representation space. This is somewhat comparable to the approach used in Principle Component Analysis, where a feature space is linearly mapped into one spanned by the principle components ordered by variance dimension. The highest eigenvalues of the feature vectors contribute to the transformation matrix based on the corresponding eigenvectors. The product of the transformation matrix with the feature vectors will then produce a linear map that is representative of the data emphasising the largest variance components and can also be used for dimensionality reduction [2] (Section 12.1).

The relevance of the CA approach for machine learning is that such a method will allow suitable representation of both static and dynamic data for classification, although the nature and number of oscillators needed to represent any given dataset is still very much under study. It is clearly the case that no arbitrary dataset is ideally suited specifically for any given network of oscillators, for which idealised classifiers can be defined, but it will allow a universal approach to the data, which is not possible without the CA representation. The CA representation permits static, dynamic and multimodal data to

be directly represented alongside each other, independently from their range or variance. As will be shown, the CA approach also allows scale-free representation and regressed representation of complex data, on which classification may readily be performed. The approach is in itself not a classifier, although it shares similarities to some clustering methods, but permits consistent and deterministic representations on which a classifier can be applied. Existing methods for regularisation focus on the reduction in outliers, such as irregular, or low-frequency occurring data values. The CA method does not seem to require this, as the resulting representation is by its nature regularised into the stable deterministic oscillation or steady state.

The relevance of this approach to biological representation should also be emphasised. The problem of any given biosystem of receiving data with high variance and logarithmic scales into a limited representation capable system of receptors or biochemical dynamic systems should not be underestimated. It has been known that a biosystem is required to be able to deal with such problems in a deterministic and reproducible manner [6]. The proposed mechanism of a network with Self-Organised Criticality properties with consistent scale-free representation based on relatively simple biochemical oscillators would allow even the smallest nontrivial biosystem to receive external input perturbation such that its own integrity and regularity are not affected, and it is able to respond in a consistent and scaled manner to those inputs.

2. Methods

The Criticality Analysis method is based on a network of Rate Control of Chaos (RCC) controlled nonlinear oscillators [3]. These oscillators are, in effect, stable oscillators due to the RCC method. However, the parameter space of the oscillators is in the chaotic domain. The oscillators are still weakly chaotic, i.e., when perturbed they will return to a periodic state, but not necessarily to the same orbit as before the perturbation due to the control. Because the RCC oscillator is deterministic, it will always go to the same orbit with the same perturbations, and the perturbation that is somewhat similar will result in a nearby, but uniquely different orbit. The network can be effectively controlled by these perturbations for this reason. It has already been shown that a network of such controlled oscillators is capable of stabilising chaotic and noisy oscillators, even when the instability is due to complexity, such as is the case in spatiotemporal chaos [3]. The RCC method adapts the local nonlinear behaviour of the system. It is conjectured that this adaptation causes small adjustments to equilibria. Such equilibrium points are moulded by the control into hyperbolic ones, if they are not already hyperbolic, and therefore acquire the associated characteristics, such as the applicability of the Hartman–Grobman Theorem [13,14] that asserts that the behaviour in a domain containing a hyperbolic equilibrium point is topologically equivalent to that of its linearization. Furthermore, a network of these RCC-controlled oscillators shows scale-free behaviour, emerging as higher-order relations between the perturbations and the total system response. This depends on the domain of control of the oscillators, and their connectivity, as would be expected [4].

In this paper, a model of a bienzymatic oscillation is used developed by Berry [15], which is described by (4) to (7). This model, which is subsequently referred to as the Berry model, has previously been shown to be controllable using the Rate Control of Chaos (RCC) method [4]. The model describes two enzymes that control the formation of extracellular matrix m from soluble filaments f . The proteinase p transforms the matrix into filaments, and conversely, transglutaminase g converts the filaments into the matrix. The extracellular matrix is also produced r_{im} at a constant rate, and each protein decays proportionally to p . r_{im} is the bifurcation parameter that causes the system to be chaotic within specific domains. By applying the Rate Control of Chaos, which is described by the quotient q_f (1), and the two control functions σ_p (2), and σ_g (3), the system remains in controlled stable orbits for a wide range of values of the bifurcation parameter. Within the subsequent simulations, the model is usually shown as a time series of the main variables m and f or as phase space

plots of f (x-axis) versus m (y-axis). Additionally, the total summed behaviour of the main variables will be referred to as M and F (capitalised).

$$q_f = \frac{f}{f + \mu_f} \quad (1)$$

$$\sigma_p(q_f) = f_p e^{(\xi_p q_f)} \quad (2)$$

$$\sigma_g(q_f) = f_g e^{(\xi_g q_f)} \quad (3)$$

$$\frac{dm}{dt} = k_g \frac{fg}{K_G + f} - \frac{mp}{1+m} + r_{im} \quad (4)$$

$$\frac{df}{dt} = -k_g \frac{fg}{K_G + f} + \frac{mp}{1+m} - \frac{fp}{1+f} \quad (5)$$

$$\frac{dp}{dt} = \sigma_p(q_f) \gamma \frac{f^n}{K_R^n + f^n} - k_a p^2 \quad (6)$$

$$\frac{dg}{dt} = \sigma_g(q_f) \beta \frac{f^l}{K_S^l + f^l} - k_{deg} \frac{gp}{K_{deg} + g} \quad (7)$$

The Berry model parameters are as follows: $\gamma = 0.026$, $\beta = 0.00075$, $K_R = 4.5$, $K_S = 1$, $K_G = 0.1$, $K_{deg} = 1.1$, $k_g = k_{deg} = 0.05$, $k_a = \frac{k_{deg}}{K_{deg}} = 0.0455$ and the Hill numbers $l = n = 4$. For different values of the bifurcation parameter r_{im} in (4), the model exhibits a wide range of dynamic behaviour, including periodic cycles, bistability and chaos [15], if left uncontrolled. The RCC control may be tuned for different amounts of control, in most cases (unless otherwise stated), the control parameters are $\mu_f = 2$, $f_p = 1$, $f_g = 1$, $\xi_p = -3$, $\xi_g = -3$. Deterministic external input ϵ is added to the r_{im} parameter as in Equation (8), and it is also used to connect the different oscillators together using a scaled relative contribution from the other $n - 1$ oscillators (without self-connections):

$$r_{im}^i = \sum_{k=1, k \neq i}^n w_k m_k + \epsilon \quad (8)$$

where w_k the connectivity strength for the oscillator which is fixed for this simulation, and importantly, remains within the chaotic domain. ϵ is a variable that represents the external data input and matches different attributes of the data provided to the input oscillators. These perturbations are fixed for a number of time steps of the evolution of the total system to allow the system to stabilise into a new orbit due to the new data presented. It should be made clear that the system will remain in the new stable orbit until the input is changed again when it will move to yet another orbit.

The Rate Control of Chaos method is furthermore used to provide dynamic scaling of the input variables. This can be employed to ensure that the input data are normalised and relies on the relative change in the input ϵ to perturb the input oscillator proportionally, which is indicated by ϵ_i for each input variable in the range $\{\epsilon_i, \dots, \epsilon_n\}$ for any n attributes of the data. This dynamic normalisation is described by the following equations.

$$q_\epsilon^i = \frac{\epsilon_i}{\epsilon_i + \mu_\epsilon^i} \quad (9)$$

$$\sigma_\epsilon^i = f_\epsilon^i e^{(\theta_\epsilon^i q_\epsilon^i q_i)} \quad (10)$$

$$r_{im}^i = \sum_{k=1, k \neq i}^n w_k m_k + \sigma_\epsilon^i \epsilon_i \quad (11)$$

where ϵ_i is the external input that represents a data point, μ_ϵ^i is the maximal value of the data (which is predetermined and constant), f_ϵ^i is a constant scalar, and θ_ϵ^i is the RCC control variable as described before, with mostly $f_\epsilon^i = 1$ and $\theta_\epsilon^i = -3$. q_i is the proportional RCC scaling for oscillator i to which the input is presented (cf. (1)); this allows the input to

be scaled in relation to the oscillator and also makes static data (such as a constant value) a dynamic variable. The RCC control is finally used to scale the input variable ϵ_i itself and is added to the r_{im} parameter of oscillator i .

The dynamic Hebbian learning rule is used for the simulations where the network is adjusted based on the cross-correlation of their activities [16]. The rule is based on the relative fraction of filament contribution for each oscillator across a connection. The weights are then described by the following equation.

$$w_{ij} = \theta_i e^{(\alpha_i q_{fi} q_{fj} \operatorname{sgn}(q_{fj} - q_{fi}))} \quad (12)$$

where the quotients q_{fi} and q_{fj} represent the proportion of filament as in Equation (1) for the respective oscillators i and j , and the weight projecting from i to j is adjusted according to (12) which strengthens the connection if the difference in these quotients is positive (i.e., the target is stronger than the source) and reduces the strength when this is not the case. Because the values of f_i and f_j may change over time, the weights may adapt to the changing behaviour of the system. Typical values for the learning parameters are $\alpha_i = -1$ and $\theta_i = 1$.

The second model used is a four-dimensional extension of the classical three-dimensional Lorenz model that incorporates an additional feedback loop in [17]. The model can exhibit chaos and hyperchaos and is described by Equations (20)–(23):

$$q_x = \frac{x}{x + \mu_x} \quad (13)$$

$$q_y = \frac{y}{y + \mu_y} \quad (14)$$

$$q_z = \frac{z}{z + \mu_z} \quad (15)$$

$$\sigma_x = e^{(\xi_x q_x q_z)} \quad (16)$$

$$\sigma_y = e^{(\xi_y q_x q_z)} \quad (17)$$

$$\sigma_z = e^{(\xi_z q_x q_y)} \quad (18)$$

$$\sigma_w = e^{(\xi_w q_y q_z + \omega)} \quad (19)$$

$$\frac{dx}{dt} = a(y - x) + e\sigma_x yz + kw + D_x \quad (20)$$

$$\frac{dy}{dt} = (cx - dy - \sigma_y xz) + D_y \quad (21)$$

$$\frac{dz}{dt} = (\sigma_z xy - bz) + D_z \quad (22)$$

$$\frac{dw}{dt} = (\rho y + f\sigma_w yz) + D_w \quad (23)$$

where $a = 56$, $b = 16$, $c = 49$, $d = 9$, $e = 30$, $f = 40$, $k = 8$, $r = -600$ for the system parameters [17] (Section 3.1), and $\mu_x = 1000$, $\mu_y = 100$, $\mu_z = 100$, $\mu_w = 3000$, $\xi_x = -1.5$, $\xi_y = -1$, $\xi_z = -1$, $\xi_w = -1.5$, $\omega = 0.055$ for the RCC parameters in the RCC Equations (13)–(19). The external input to the oscillators is provided by $\rho = -100\epsilon$ with ϵ as the perturbation value representing data for each oscillator. Additionally, the n oscillators are connected to each other using the diffusion terms D_x , D_y , D_z , D_w where each term is simply the weighted sum of the corresponding variables for the oscillators, e.g., $D_x = \sum_{k=1}^n w_k x_k$, with optional weighting term w_k . Note that the bifurcation parameter r is chosen in the hyperchaotic domain; see Figure 1 in [17].

The models were simulated using the EuNeurone, version 2.3 numerical integration software, which is available on Zenodo [18]. The results can nevertheless be readily simulated using other fixed-step numerical integration tools. The numerical integrators used were the standard fixed step integration Runge–Kutta RK4, and Fehlberg RK algorithms [19] (pages 363–366). Results were then exported into Hierarchical Data Format 5 (HDF5), subsequently analysed, and plotted using Matlab.

The datasets used in these experiments are the standard, unmodified versions of the classic Iris dataset (3 classes, 4 attributes, 120 samples, numerical), the Wine dataset (3 classes, 13 attributes, 178 samples, multivariate), and the Bonemarrow dataset (unclassified, 36 attributes, 172 samples) from the UCI Machine Learning Repository [20]. The order of the samples was randomised before input to the oscillators, and the category labels were not presented to the networks; these were only used to determine the class after the simulations. With the Bonemarrow dataset, all incomplete samples were excluded, and three of the attributes were unused. The main aim of this dataset is to test for possible relations between some attributes and survival [21]. It is important to note that the data in these models are used to generate the nonlinear representation spaces, not primarily to provide categorisation directly. These models are shown to allow a biologically relevant mechanism to represent dynamically changing, as well as static, data to be represented in a manner that facilitates categorisation and is insensitive to scale, but sensitive to variability within the sample dataset.

3. Results

Let us first consider the simplest representation case, where the four-dimensional Iris dataset is projected onto a small network of four Berry oscillators. We will then explore the effect of Hebbian learning, reservoir computing, normalisation and the persistence of the generated representations.

3.1. Simply Perturbed Berry Model

The four oscillators are RCC controlled, as previously described, and have mutual weights of $w_k = 0.0005$, and the Iris data is scaled by 0.00025 when added to Equation (8) as the parameter ϵ . Figure 1A shows the first twelve data points of the Iris datasets as the four values presented to the network. Each data point is clamped to the input for 50,000 evolution time steps where the step size is 0.1. In the next panel, Figure 1B shows the dynamic response of the network of those twelve data points, as the total amount of M of the four oscillators, colour-coded to the correct class. The class type itself is not used by the network as such. Panel C shows the resulting dynamic response due to the data perturbation of the entire Iris dataset as a phase space representation of the total F versus M . Here, the blue class (*Iris setosa*) causes the largest perturbations, followed by the red class (*Iris virginica*) and finally the green class (*Iris versicolour*). Note that the last two are the classes that require nonlinear separation boundaries to classify [20]. To allow for readily classification, although the aim here is to create dynamic representation spaces, not to optimise for classification, the maxima of each of the total oscillations are plotted in panel D. It should be clear that the representation is not perfectly separable given the data, but that even a simple perceptron classifier can be readily trained to perform reasonably well on this representation using two simple linear boundaries. This representation of the Iris data is due to the dynamic response of the network to the perturbations that represent the data points and is emphasised by the criticality properties of the controlled chaotic system. Additionally, this dynamic response requires only relevant input, and no training for the representation is necessary.

To demonstrate this interesting property of criticality in response to deterministic perturbations, the network is extended to 64 RCC-controlled Berry oscillators. The Iris data are presented to the network as input to only the first four oscillators with a scaling of 0.005, and the network weights are reduced to $w_k = 0.00005$. Apart from the reduction in overall connectivity strength and the number of oscillators, the model remains the same. Due to

the critical response, i.e., the ability to change state from one orbit to another due to small perturbations (in this case the data), the network response is qualitatively similar to the small network of four oscillators, as can be seen in Figure 1E, where the phase space plot of the orbits colour coded to the categories is shown. Subsequently, Figure 1F shows the maxima of these orbits on which a classifier can readily be trained.

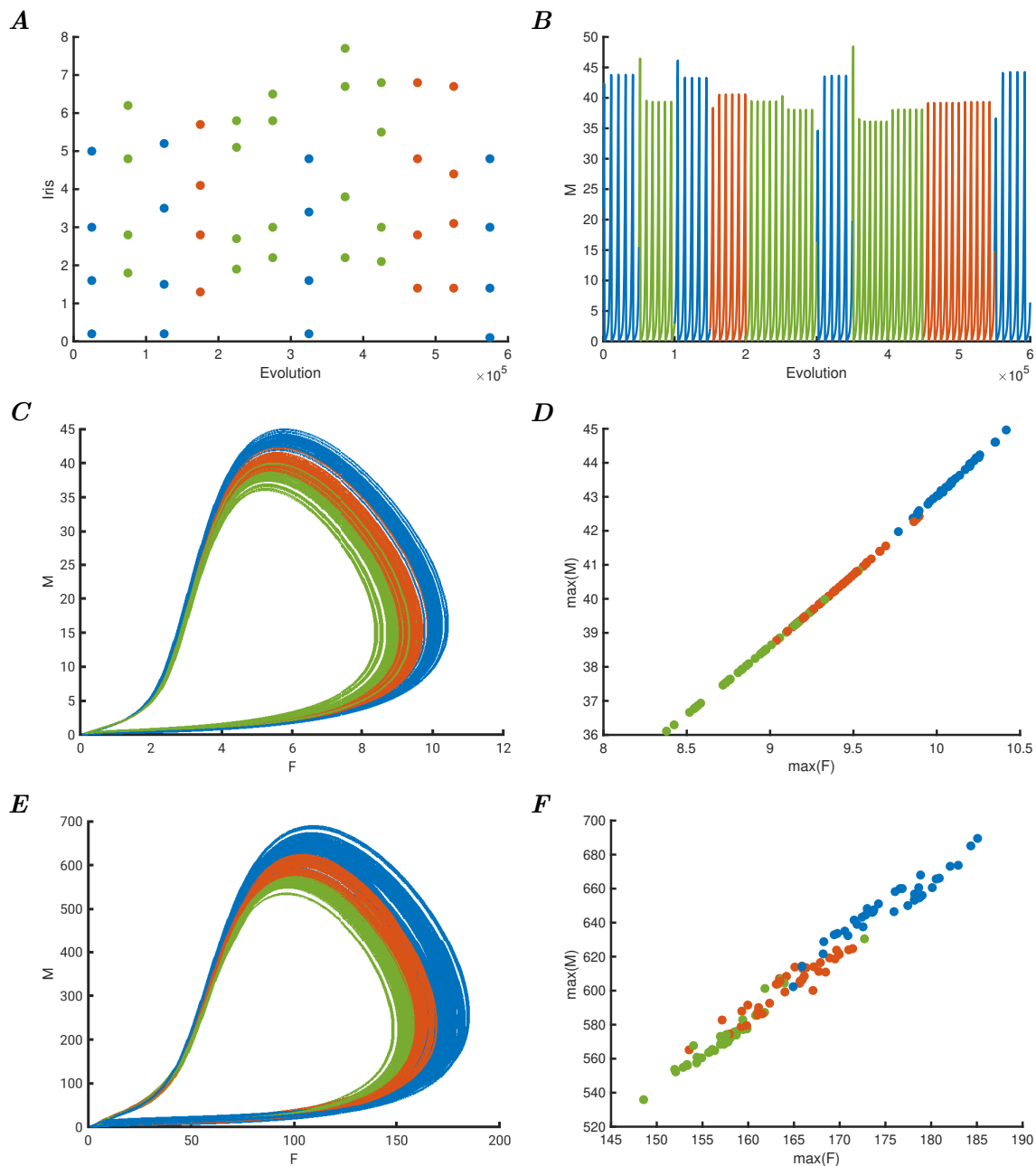


Figure 1. (A) First 12 samples of the Iris dataset as presented to the 4-oscillator network. (B) Total M resulting from the perturbations caused by the Iris dataset of the network. (C) Total F versus M of the 4-oscillator Iris perturbed network showing different dynamic domains representing categories. (D) Maxima of F versus M of the figure in panel (C) demonstrating possible linear separation boundaries. (E) Iris perturbed phase space plot of F versus M of the 64-oscillator network, showing size-independent constant representation of the data. (F) Phase space plot of the maxima of the 64-oscillator Iris perturbed network, showing consistent representation of the data.

3.2. Oscillators with Dynamic Hebbian Learning

It is well understood that the connectivity weights in a network of oscillators can affect the behaviour of the system, and to explore this property, a subsequent set of experiments was devised. A network of four oscillators is adapted with the dynamic Hebbian learning rule according to Equation (12). The weights respond to the changing behaviour of each oscillator from i to j where the latter is the target and the former is the source. The external data input is provided as in the previous models, but with enhanced input strength of 0.005 as described before. The learning parameters are $\alpha_i = -1$ and $\theta_i = 1$ for all connections in this experiment. The results are shown in Figure 2A, where twelve orbits of the total M are shown over time. Panel B shows the phase space plot of F versus M of the entire Iris dataset. In panel C, the maxima of these two variables are shown. At first glance, it may appear that the representation in this manner makes it harder to classify; however, as can be seen from the first panel in Figure 2A, the shape of the oscillation has changed rather than the amplitudes. This shows that the connectivity of the critical network of oscillators affects the dynamic behaviour of the system, rather than the individual amplitudes or peaks of each of the periods. To classify such dynamic behaviour, in a meaningful biologically relevant sense, would require the subsequent classifier to be able to receive the dynamic input of the representation and separate them as different evolutions over time. To illustrate this further, the median frequency of each total oscillation for the total value of F was estimated and plotted against the maxima of the total value of M . This is shown in panel D, where the adjusted mean frequency estimate can be used in combination with the maxima to separate the classes, in a non-optimal but suitable manner. The importance is not the efficacy of the classifier, but the ability to represent complex data in different ways that are meaningful for continuous dynamic systems such as biosystems to permit classification in the first place.

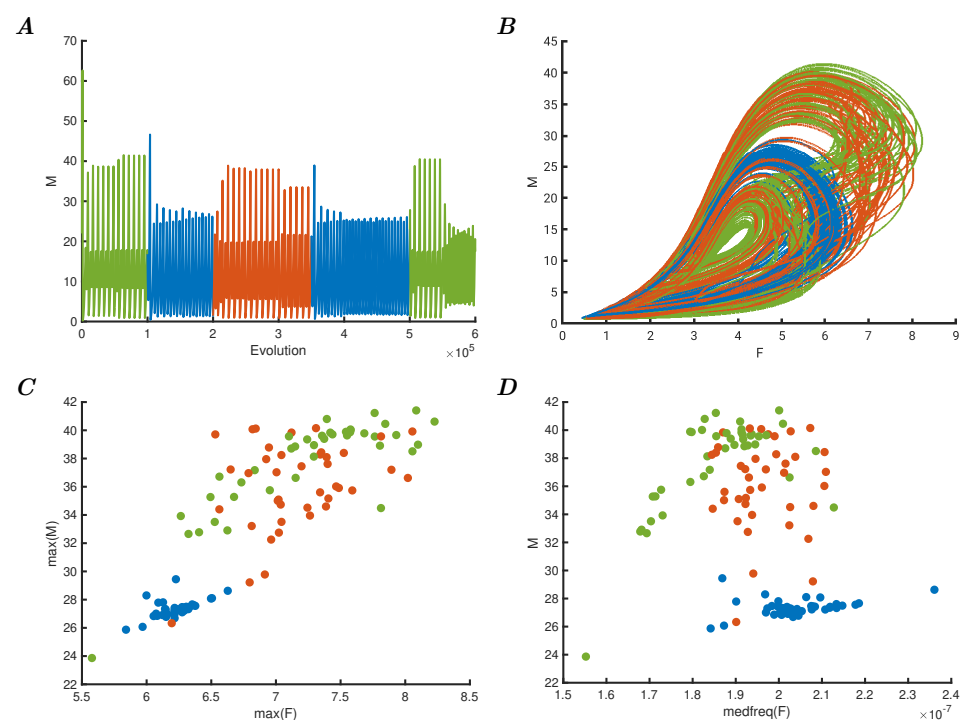


Figure 2. (A) Evolution over time of 12 samples from the Iris dataset input to a network of oscillators with dynamic Hebbian learning. (B) Phase space plot of the total F versus M of all the Iris samples, showing changes in dynamic response due to the perturbations causing adaptive learning. (C) Maxima of total F versus M of all the Iris samples in the network, showing poor separation on amplitude. (D) Median frequency of the total F versus the maximal M showing that separation is possible due to the changing dynamics of the orbits, as is shown in panel (A).

3.3. Deterministic Reservoir Computing

The property of changing dynamics based on external perturbation may be exploited in a different manner, as part of a representative state based on these perturbations. Such a reservoir of dynamic behaviour requires a readout layer for categorisation. To explore this aspect, another network was devised that contained six Berry-type oscillators with two additional external oscillators that may act as readout units. These two external oscillators are configured similarly to the six reservoir units. In this simulation, the network of six oscillators that form the reservoir does not contain any learning, similar to the four units in the first model, but the learning rule is only applied to the connections of all six units to the readout units [10]. The data are provided to four of the six reservoir units in Equation (8). The input to readout units 7 and 8 is given by

$$r_{im}^7 = \phi_7 \sum_{k=1}^{n=6} w_k F + m_8 \quad (24)$$

$$r_{im}^8 = \phi_8 \sum_{k=1}^{n=6} w_k M + m_7 \quad (25)$$

where F and M are the summed values of the six f and m variables, respectively, and $\phi_7 = \phi_8 = 0.00015$ as scalars. w_k represents the weight from each of the six units to unit 7 and 8 in each equation and is determined by the dynamic Hebbian learning rule (12). In other words, the input to each unit here is the summed weight in proportion to the total value of F of the reservoir for unit 7, and in proportion to the total value of M for 8. In addition, the value of m_8 is added to unit 7 and vice versa to reflect the effect the units have on each other. The parameter values are the same as in the previous models, except that the external input is scaled to 0.00225, and the fixed connections in the reservoir are 0.0002. The learning parameters are $\alpha_i = -3$ and $\theta_i = 3$ for the two readout units.

The result of presenting this reservoir with the Iris dataset is shown in Figure 3. Panels A to C show five arbitrarily chosen dynamic responses to each of the three categories as observed by readout unit 7 over time. Because the connectivity strength is higher than in the first experiment, the network responds with different oscillations due to the perturbations of each data point. As can be seen, there are significant differences in these responses, and although some points are not easily defined, they are each recognisably representative of their class. Panel D shows the maxima of all these oscillations for the Iris dataset, which can be compared to the single orbit oscillations that characterised the first model (described in Figure 1). Furthermore, by comparing the individual readout units 7 and 8 in Figure 4, it is clear that these units independently show different aspects of the reservoir behaviour. Panel A shows the maxima of unit 7, and C the maxima of unit 8. The first shows the same characteristic behaviour as the overall reservoir (compared to Figure 3D), and panel D shows the characteristic behaviour as comparable to the original model in Figure 1D. It can therefore be concluded that the readout units show a dynamically reduced but characteristic behaviour of the entire reservoir, i.e., each individual oscillator shows the same behaviour for which the entire network is required to produce. The exact nature of the properties of the classifier required to actually separate these classes using either of the readout units (or both) is not essential for this argument, but it is clearly much easier to classify on one oscillator's behaviour rather than a network, or indeed the data itself. To complete the comparison of this reservoir computing approach, the median frequencies of the two readout oscillators are also determined and shown in Figure 4B for unit 7 and in D for unit 8, showing that the median frequency versus the maxima of f_7 and f_8 , respectively. These are different again from the previous model experiment (in Figure 2) due to the lack of adaptation in connections and the overall readout of the network.

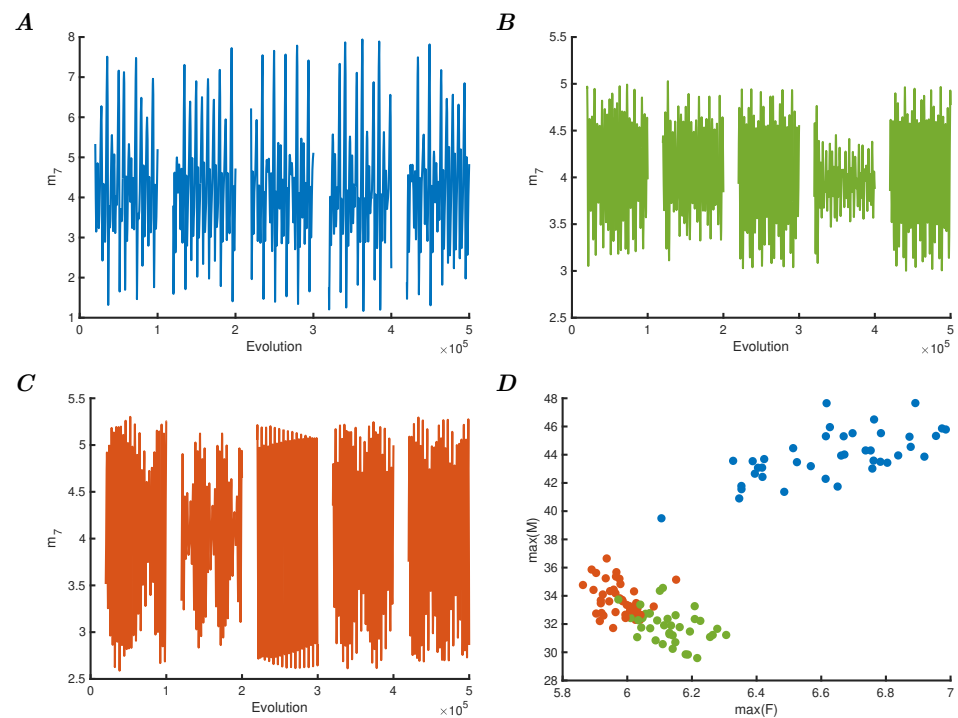


Figure 3. (A) Orbits of 5 samples of the first class in the Iris dataset as represented by the readout unit 7 over time. (B) Orbits of 5 samples of the second class in the Iris dataset as represented by unit 7 over time. (C) Same as previous panels, but for the third class. (D) Phase space plot of the maxima of total F and M of the six reservoir units (i.e., excluding the readout units 7 and 8).

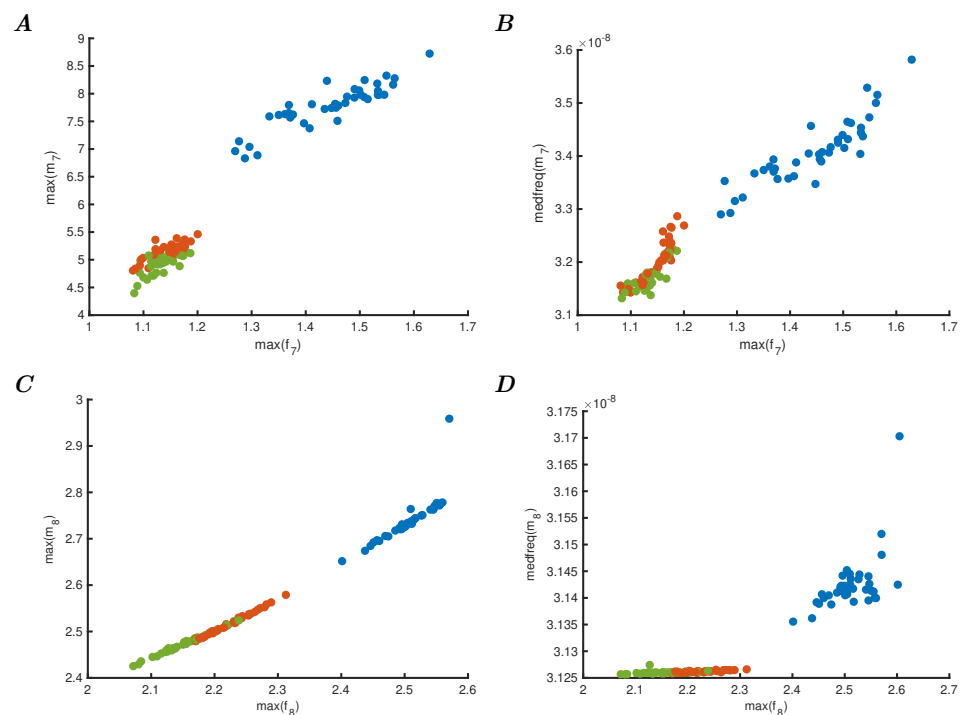


Figure 4. (A) Phase space plot of the maximal f and m from the readout unit 7. (B) Phase space plot of the maximal f versus the median frequency of m of unit 7. (C) Phase space plot of the maximal f and m of the readout unit 8. (D) Phase space plot of unit 8, showing maximal f versus the median frequency of m .

3.4. Simply Perturbed Wu Model

The previous experiments may give the impression that the choice of the Berry model is essential to these results, but apart from the requirement of the RCC control to provide the critical state, different models can be used. As an illustration, the Iris data were presented to a network of Wu models. These oscillators are an extended version of the classical Lorenz model. The parameters are described in the Methods section, and the results can be seen in Figure 5. Panel A shows W (the summed value of the four w variables of the four models) over time for 20 samples, colour-coded for each category. Notice that the system is now not oscillating, but moves to representative controlled steady states. Panel B shows the phase space plot of the maxima of Z versus W (after transients have been removed), in effect showing the steady state points. The relative location of these points is comparable to the ones produced by the Berry model due to the fact that the perturbations cause somewhat similar dynamic responses, even if the quantitative response is very different. That this is the case for a specific perturbation and made visible for a specific projection of the representation space can be seen in panel C, where the phase space plot of the maxima of three of the four system variables is shown, with the relative representation clearly visible. The simple perturbed model approach is appropriate for the Iris dataset and would allow readily categorisation based on the network response. It is worth noting here that nonlinear models may be devised to fit arbitrary limit cycle shapes [22], but perturbing networks of such models would not provide similar results due to the SOC properties of the RCC-controlled network not present in the former network of models.

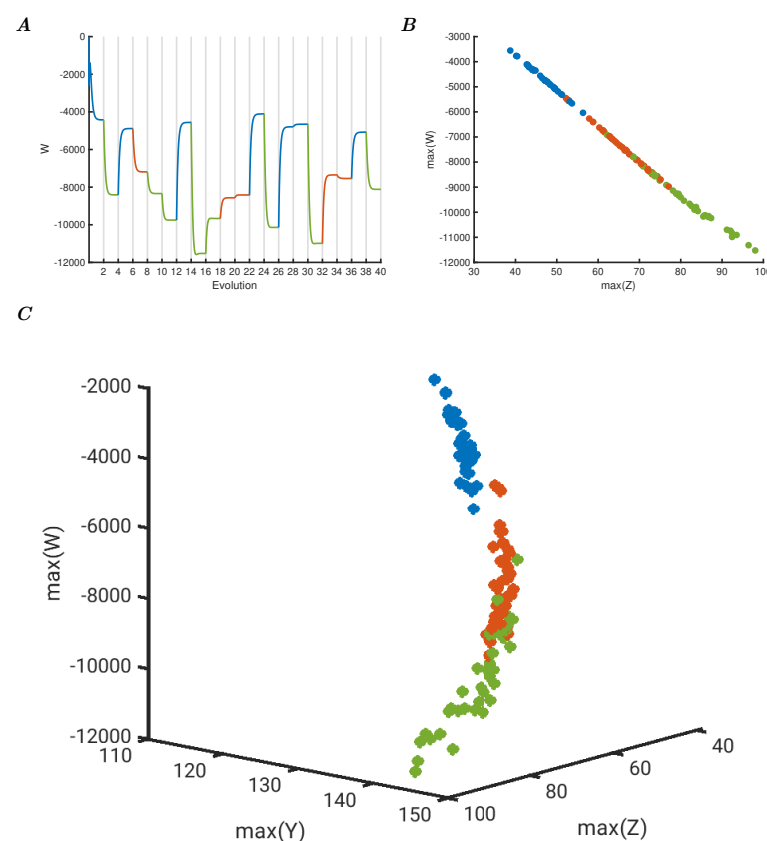


Figure 5. (A) Network of four Wu oscillators with Iris data input, showing the evolution of 20 samples with total W of the network. The network does not oscillate and converges to steady states, as shown in the plot where the transients end at a specific value of total W . (B) Phase space representation of the maximal total Z versus W showing the steady state representations in two dimensions allowing separation. (C) Three-dimensional representation of the maxima of total Y , Z and W , with the Iris data presented to the four Wu oscillator network, showing different representation in multiple dimensions.

3.5. Dynamic Normalisation

Normalisation is often used to ensure that the data with large variations is within the boundaries of the representation space. It allows the data to be scaled and represented as uniformly as possible for the training of a classifier. Here it is shown that the use of a dynamic normalisation method has the same effect as it would have for a normal n -dimensional representation space. The standard Wine dataset with 3 classes and 13 attributes is used to show that normalisation, using the method described by Equations (9)–(11), improves the ability of the network to create a nonlinear representation space for subsequent classification. This dataset has large variations between attributes, and this can skew the ability of a classifier, which is why it tends to be normalised before processing anyway. Figure 6 shows the input of Wine data into the 64 Berry oscillator network. Panel A and B show the data as unscaled input, i.e., as is, with scalars of 0.001 for the first 12 attributes, and 0.0001 for the last attribute, which is about two orders of magnitude larger than the other attributes. Additionally, the network connectivity is initially randomised around a mean of 0.000035 with a variance of 0.00001, and it is subsequently kept at the same value. This shows that the network connectivity changes the behaviour but not the dynamic response to individual data attributes. The figures show that the system behaves similarly to the Iris data (that has only four attributes), but is not readily classifiable on the amplitudes. Using the dynamic normalisation on top of the same approach (as described in the Methods section), but with everything else kept the same (including initial random connectivity strengths), the result shows that now the network can represent these data points as nonlinear orbits (panel C), whose maxima could conceivably be classified (panel D).

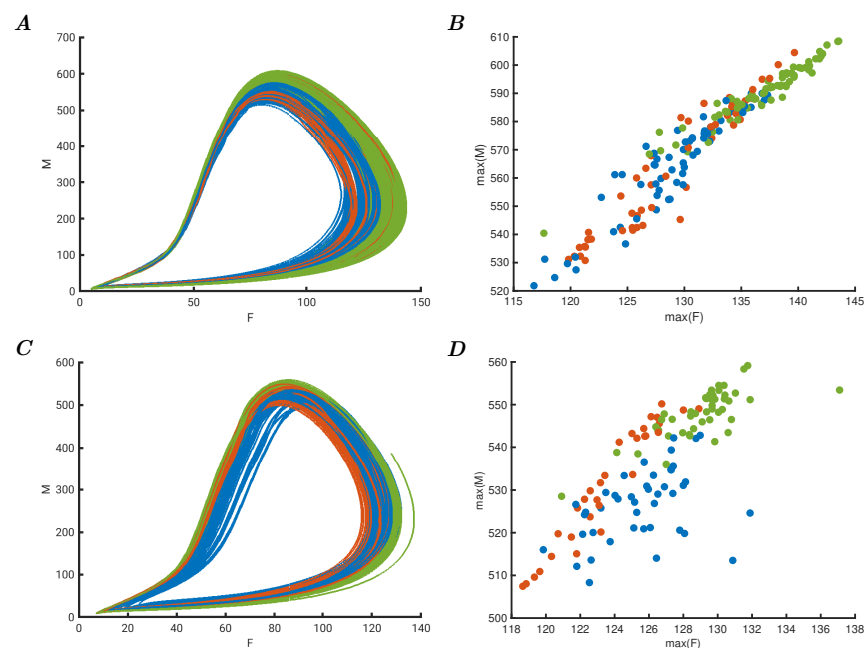


Figure 6. (A) Network of 64 Berry oscillators with unmodified, scaled input from the 13-attribute Wine data showing phase space of all samples as total F versus M . (B) Maximal F versus M of all the Wine samples, showing poor inherent separation based on amplitude. (C) Same 64-oscillator network, but with a dynamic normalisation function to adjust the input representation to the network, showing the phase space of total F versus M . Note that the position of the samples in the nonlinear representation has changed. (D) Maximal F versus M of the dynamic normalised Wine data of all samples, showing that the normalised input has improved the dynamic nonlinear representation such that some separation will be possible.

3.6. Representational Persistence

After addressing the nonlinear representation available using Criticality Analysis for classification tasks, the manner in which CA behaves with respect to unknown relations within the data needs to be considered. To this end, a simulation was performed using a non-categorical dataset to demonstrate that the resulting dynamic behaviour is still consistent with the input and that the representation of known shared feature relations persists in the CA representation.

Figure 7 shows the nonlinear Criticality Analysis representation of 142 individuals from the Bonemarrow dataset [21]. This was made using the 64 Berry oscillator network with normalisation of the 34 input features using scalars of 0.0001 and network connectivity strength of 0.00005. The dataset is not specifically designed for classification or regression, but to determine if some of the features increase the likelihood of survival or quality of health. Specifically, the increased dosage of CD34+ cells per kg may extend survival. The survival status is shown in panel A, where no clear pattern emerges, which is consistent with the original results. The representation should show if any of the features are similar to shared features within the dataset. If the feature is persistently responsible for a nonlinear representation state, this would show as a grouping within the nonlinear representation. In the absence of such recognisable groups, or the inability to define such groups with possible separation boundaries, the data may not contain the required information to allow this, or there is too much variability (possibly due to noise). In this case, comparing the features with the representation shows that none of the features show any relation to the categories that these represent, apart from the two age-related features. It is known that both the age of the donor and the age of the recipient are factors that exist persistently within the data representation [21]. Therefore, the resulting Criticality Analysis representation should show these relations. Panel C shows the representation of the donor age showing clustering between the two classes (donors younger than 35 years, and older). Additionally, panel D shows the two age classes associated with recipients of the two age groups (recipients younger than 10 years, and older). The age relations are categorisable and demonstrate that their relations persist within the CA representation even though the other features do not.

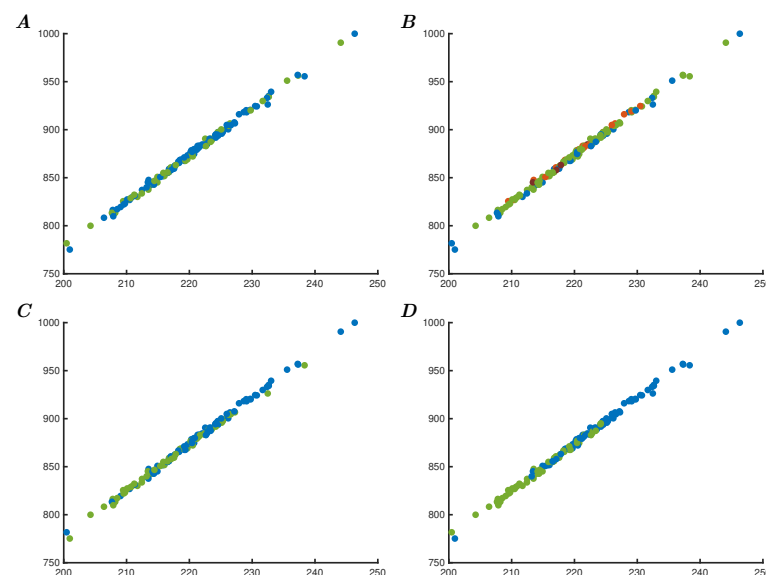


Figure 7. (A) A network of normalised input of the Bonemarrow data colour-coded to survival of 64 Berry oscillators showing little relation between the two categories. (B) Same network coded to the aimed relation to the relative increase in CD34+ cells/kg showing also little relations between individuals. (C) Same network coded to the known relation with the bone marrow donor age (blue < 35, green ≥ 35 years old). (D) Network with colour coding to the also known relation with bone marrow recipients age (blue < 10, green ≥ 10 years old). The last two figures show clear grouping expressing the persistence of the age relation within the CA nonlinear representation.

4. Discussion

One of the issues with categorisation problems is that the representation of data should be in a manner such that the categorisation method requires minimal effort to train. Commonly, a large amount of effort is required to optimise this process. For many datasets, this is not readily possible, due to the temporal and dynamic nature of the data, and the static nature of the learning process. Although highly optimised algorithms can be devised for specific datasets, and some significant improvement has been made to allow forms of adaptive learning, these methods are not generally applicable to dynamic data.

Biosystems need to solve problems at every level of existence, from local low-level biochemical behaviour to high-level human thought. The overall idea of nonlinear critical data representation is that it allows different levels to function reliably. The method of Criticality Analysis aims to provide a much-reduced mechanism for data representation that is robust and biologically relevant, as it is based on simple oscillators that can exhibit a consistent and deterministic representation of the input. The main requirement is the applicability of Rate Control of Chaos on a dynamic nonlinear system such that it can exhibit criticality. Here, criticality has been defined as the property of a dynamic system to change state due to small perturbations. When these perturbations are deterministic, as they represent data points, the result should therefore be deterministic behaviour as well, as can be seen in the results presented previously.

It should be clear that the network of dynamic oscillators does not require training to create any representation. There is some need for tuning, where a minimum level of connectivity is required to ensure that the total behaviour is critical, and there is a maximal connectivity level that would destabilise the network's behaviour. This level depends on the total input of multiple oscillators, and the scalar used for the input data, which were determined experimentally. However, as can be seen in Figures 2–4, training may allow the creation of representations on which a categorisation algorithm can more readily be trained. This would permit an arbitrary dataset to be presented to the network, and then a given learning algorithm may be able to cause the best representation to achieve categorisation. This does not necessarily lead to optimal representations or best categorisation, but as can be seen from the examples shown of the networks based on the Berry and Wu models, a straightforward representation can produce acceptable results with little effort.

Furthermore, the training of the reservoir can also allow the reduced representation of a more complex dataset. The algorithm used for training could readily be based on backpropagation or a similar standard neural network approach. It can therefore be argued that Criticality Analysis, if based on an adaptive suitable set of oscillators, would allow the learning of complex data to become possible even for a relatively simple organism, just by adjusting the weighting of the input of biochemical oscillators.

The data provided to the network need to be scaled to ensure that the total input is not destabilising the network, in a similar manner as the connectivity. The input data act as the perturbation to the other oscillators, which explains the deterministic nature of the perturbation, as each acts as a constant input to the RCC that stabilises the critical system within each individual controlled oscillator. If the input is too strong or too weak, the network will simply not be in a critical state. The use of the normalisation approach may help here, but possibly at the price of a loss of information that may exist in the relative differences between attributes. The performance of the CA approach depends on the number of oscillators used to represent the data, the numerical complexity of each oscillator, and the duration of each perturbation when presenting the network with a data sample for the system to stabilise into a steady state or orbit. For example, a more biological extensive representation of a chemical oscillator may require more numerical integration steps for each oscillator to calculate its evolution, which then needs to be calculated for every oscillator in the network. A parallel computational implementation of the software is under development and will improve performance.

Features that are persistent within the dataset are preserved within the Criticality Analysis representation. The existence of hidden or convoluted correlations is therefore not affected using the method, although it does not aid in identifying those especially.

To demonstrate the utility of the CA method for dynamic biological data, we have recently shown in a proof of concept experiment using IMU sensors that the method can be used to recognise impaired gait from the raw gait data alone [23]. Subsequently, we extend this proof to a large gait dataset of healthy normal volunteers to show correct high-performance classification using CA with SVM classifiers. Additionally, we have measured gait in children has been used in combination with CA to assess improvements in health due to clinical treatment [24].

Lastly, the method described in this paper may be considered to be some form of a clustering algorithm. However, that would neglect the uniqueness of the representation, and the deterministic nature of the method. The relation between the data properties and the critical representation is still under investigation. It should be recognised that not any given dataset can be best represented by any given critical model, in which case, only one model would fit all datasets. Determining which is the most critical model for a specific data representation, as well as the development of further models that have suitable properties for this type of analysis, is still underway.

5. Conclusions

The results show that the nonlinear representation using Criticality Analysis does not by itself represent underlying correlations, but allows arbitrary data to be categorised using an appropriate set of oscillators with a suitable categorisation method. It should be clear that there is no fundamental reason that these oscillators are specifically useful to represent the attached data, nor are they especially suited for these datasets. The aim is to show that it is firstly possible to represent data in a dynamic and biologically relevant manner, which can also be used as a reservoir computing approach, that allows dimensional reduction without loss of representational behaviour. Secondly, the data can be categorised without any learning in the best cases in which a simple categoriser (such as a perceptron) can readily be trained. Thirdly, dynamically changing data are the norm within biological systems, and the proposed mechanism may be one of the ways in which biological data may be represented in a size-independent but consistent manner due to the critical nature of the network of oscillators. The Criticality Analysis method can be used to create a suitable representation of complex, dynamic data for further optimised categorisation approaches, but may also shine a light on one of the most fundamental problems in biosystems, that of reliable, deterministic, and scale-free data representation.

Funding: This research received no external funding.

Data Availability Statement: For each figure in the manuscript, a corresponding model file is available from Zenodo [25] that allow reconstruction of the results.

Acknowledgments: Author is grateful to Oxford Brookes University for providing support towards this publication.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CA	Criticality Analysis
ESN	Echo State Network
IMU	Inertia Measurement Unit
LSM	Liquid State Machine
PCA	Principle Component Analysis
RCC	Rate Control of Chaos
SOC	Self-Organised Criticality
SVM	Support Vector Machine

References

- Ribeiro, M.T.; Singh, S.; Guestrin, C. Why Should I Trust You? Explaining the Predictions of Any Classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144.
- Bishop, C. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.
- olde Scheper, T.V. Biologically Inspired Rate Control of Chaos. *Chaos Interdiscip. J. Nonlinear Sci.* **2017**, *27*, 103122. [[CrossRef](#)] [[PubMed](#)]
- olde Scheper, T.V. Controlled Bio-inspired Self-Organised Criticality. *PLoS ONE* **2022**, *17*, e0260016. [[CrossRef](#)] [[PubMed](#)]
- Mora, T.; Bialek, W. Are Biological Systems Poised at Criticality? *J. Stat. Phys.* **2011**, *144*, 268–302. [[CrossRef](#)]
- Kaufman, S. *The Origins of Order: Self Organisation and Selection in Evolution*; Oxford University Press: Oxford, UK, 1993.
- Watkins, N.W.; Pruessner, G.; Chapman, S.C.; Crosby, N.B.; Jensen, H.J. 25 Years of Self-organized Criticality: Concepts and Controversies. *Space Sci. Rev.* **2016**, *198*, 3–44. [[CrossRef](#)]
- Natschläger, T.; Maass, W.; Markram, H. The “Liquid Computer”: A Novel Strategy for Real-Time Computing on Time Series. *Telematik* **2002**, *8*, 39–43.
- Jaeger, H. Adaptive nonlinear system identification with echo state networks. In Proceedings of the Neural Information Processing Systems, Vancouver, BC, Canada, 9–14 December 2002.
- Steil, J.J. Backpropagation-Decorrelation: Online recurrent learning with O(N) complexity. In Proceedings of the IEEE International Conference on Neural Networks, Budapest, Hungary, 25–29 July 2004; Volume 2, pp. 843–848. [[CrossRef](#)]
- Dutoit, X.; Schrauwen, B.; Campenhout, J.V.; Stroobandt, D.; Brussel, H.V.; Nuttin, M. Pruning and regularization in reservoir computing. *Neurocomputing* **2008**, *72*, 1534–1546. [[CrossRef](#)]
- Livi, L.; Bianchi, F.M.; Alippi, C. Determination of the Edge of Criticality in Echo State Networks Through Fisher Information Maximization. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 706–717. [[CrossRef](#)] [[PubMed](#)]
- Hartman, P. A lemma in the theory of structural stability of differential equations. *Proc. Am. Math. Soc.* **1960**, *11*, 610–620. [[CrossRef](#)]
- Grobman, D.M. Homeomorphism of systems of differential equations. *Dokl. Akad. Nauk SSSR* **1959**, *128*, 880–881.
- Berry, H. Chaos in a bienzymatic cyclic model with two autocatalytic loops. *Chaos Solitons Fractals* **2003**, *18*, 1001–1014. [[CrossRef](#)]
- Brown, R.E. Donald O. Hebb and the Organization of Behavior: 17 years in the writing. *Mol. Brain* **2020**, *13*, 55. [[CrossRef](#)] [[PubMed](#)]
- Wu, W.; Chen, Z.; Yuan, Z. The evolution of a novel four-dimensional autonomous system: Among 3-torus, limit cycle, 2-torus, chaos and hyperchaos. *Chaos Solitons Fractals* **2009**, *39*, 2340–2356. [[CrossRef](#)]
- Zenodo. *EuNeurone v2.3*; Zenodo: Geneva, Switzerland, 2013. [[CrossRef](#)]
- Gough, B. *GNU Scientific Library Reference Manual*, 3rd ed.; Network Theory Ltd.: Bristol, UK, 2009.
- Dua, D.; Graff, C. UCI Machine Learning Repository. 2017. Available online: <http://archive.ics.uci.edu/ml> (accessed on 11 December 2012).
- Sikora, M.; Wróbel, L.; Gudyś, A. GuideR: A guided separate-and-conquer rule learning in classification, regression, and survival settings. *Knowl.-Based Syst.* **2019**, *173*, 1–14. [[CrossRef](#)]
- Ajalloeian, M.; Kieboom, J.V.D.; Mukovskiy, A.; Giese, M.a.; Ijspeert, A.J. A general family of morphed nonlinear phase oscillators with arbitrary limit cycle shape. *Phys. D Nonlinear Phenom.* **2013**, *263*, 41–56. [[CrossRef](#)]
- Eltanani, S.; olde Scheper, T.V.; Dawes, H. A Novel Criticality Analysis Technique for Detecting Dynamic Disturbances in Human Gait. *Computers* **2022**, *11*, 120. [[CrossRef](#)]
- Eltanani, S.; Scheper, T.V.O.; Munoz-Balbontin, M.; Aldea, A.; Cossington, J.; Lawrie, S.; Villalpando-Carrion, S.; Adame, M.J.; Felgueres, D.; Martin, C.; et al. A Novel Criticality Analysis Method for Assessing Obesity Treatment Efficacy. *Appl. Sci.* **2023**, under review. [[CrossRef](#)]
- Zenodo. *Criticality Analysis Equation Files*; Zenodo: Geneva, Switzerland, 2022. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.