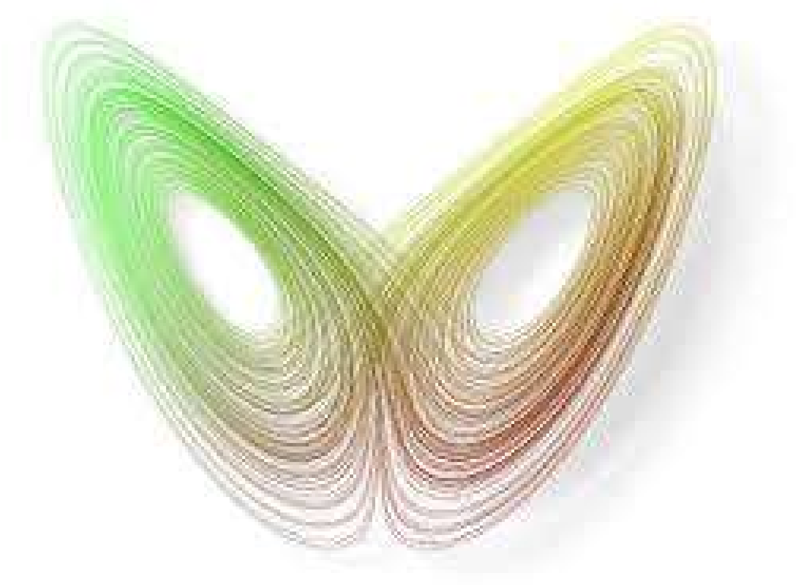
Machine learning and empirical dynamic modeling to study non-linear dynamic systems



In the phase portraits of many systems, there are fixed points toward which the system evolves.

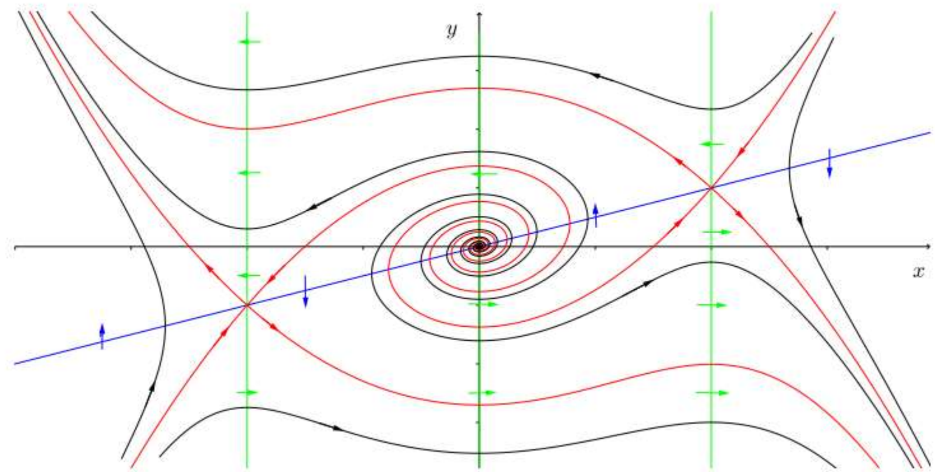


Figure 14: Phase portrait of system (43). Note the horizontal isoclines are coloured green, while the vertical isocline is coloured blue.

Image from https://math.stackexchange.com/questions/1275448/dynamical-systems-plotting-phase-portrait

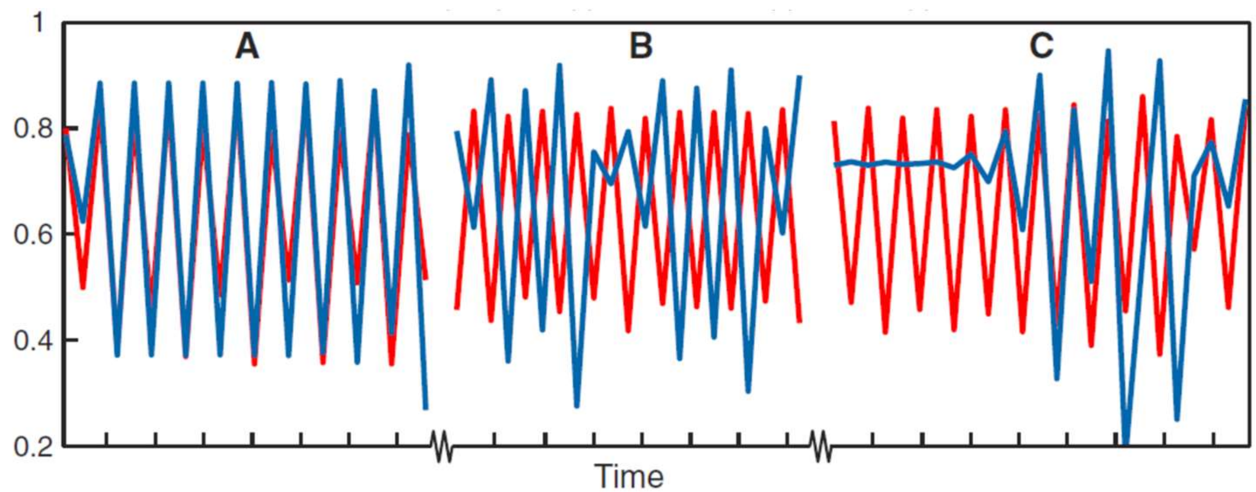
Deep Climate NYC, 9 October 2018

Sylvia Sullivan

In nonlinear dynamical systems,

correlation does not imply causation,

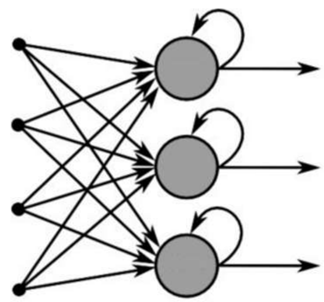
and lack of correlation does not imply lack of causation.



So we need more sophisticated techniques to tease out causality or forecasting ability.

Image from G. Sugihara et al. *Detecting causality in complex ecosystems.* Science **338** 496-500 (2012).

The idea of a dynamical system is relevant for a machine learning model if it exhibits **recurring behaviors**.



Recurrent Neural Network

are useful for **sequential data**
and longer-term dependencies.

$$\dot{x}_1 = (1 - x_1^2)x_2,$$
$$\dot{x}_2 = x_1/2 - x_2.$$

translate this dynamical system to an RNN

activations recurrences

$$\dot{x}_i = -x_i + \sum_k^N J_{ik} r_k + \sum_k^I B_{ik} u_k,$$

weights

$$r_i = h(x_i),$$

Hidden units that affect themselves
as well as on to the output units.

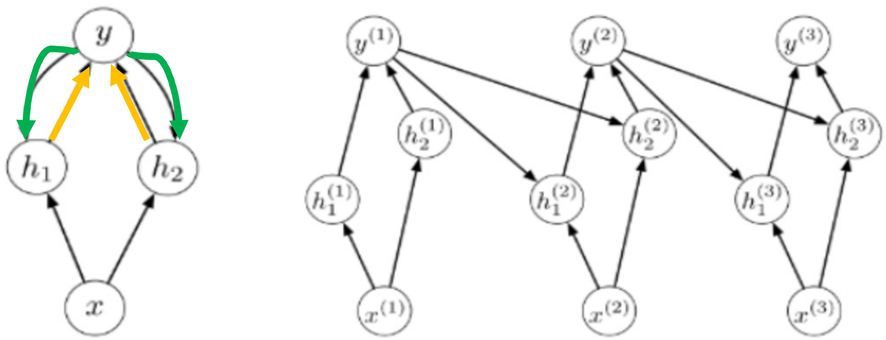
D. Sussillo and O. Barak. *Opening the Black Box: Low-dimensional dynamics in high-dimensional recurrent neural networks.* Neural Computation **25** 626-649 (2013).

Parity is a simple example.

Inputs = 0 1 0 1 1 0 1 0 1 1 …

Output = 0 1 1 0 1 1 0 0 1 0 …

We use the input now and the output previously to calculate our hidden weights.



activations recurrences

$$\dot{x}_i = -x_i + \sum_k^N J_{ik} r_k + \sum_k^I B_{ik} u_k,$$

weights

$$r_i = h\left(x_i\right),$$

Ideas taken from http://www.cs.toronto.edu/~rgrosse/courses/csc321_2017/slides/lec14.pdf

How does a recurrent neural network operate?

In the phase portraits of many systems, there are fixed points toward which the system evolves.

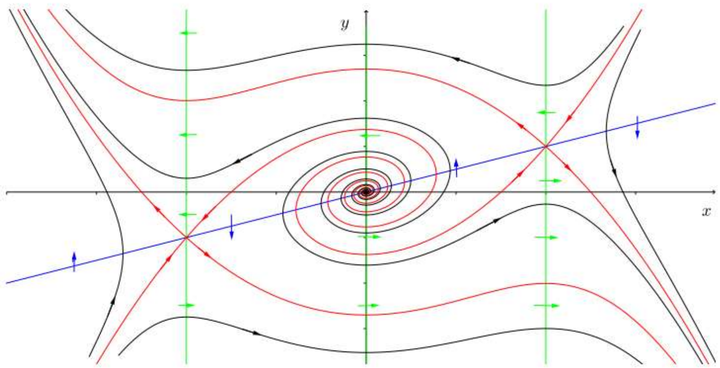
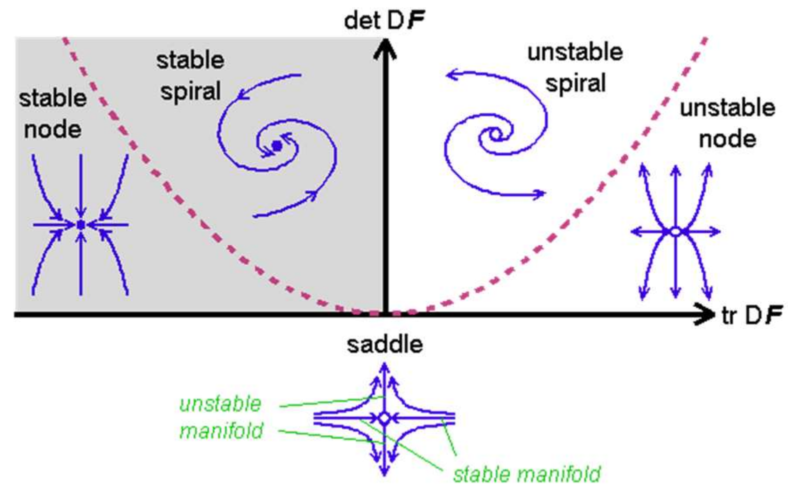


Figure 14: Phase portrait of system (43). Note the horizontal isoclines are coloured green, while the vertical isocline is coloured blue.



Linearization around fixed points (or slow points) in the phase space of the dynamical system gives insight into how the RNN works.

Righthand image from S. H. Strogatz. *Nonlinear Dynamics and Chaos with applications to Physics, Biology, Chemistry, and Engineering.*

D. Sussillo and O. Barak. *Opening the Black Box: Low-dimensional dynamics in high-dimensional recurrent neural networks.* Neural Computation **25** 626-649 (2013).

A sine wave generator is another simple example of how RNNs employ linearized dynamics about fixed points.

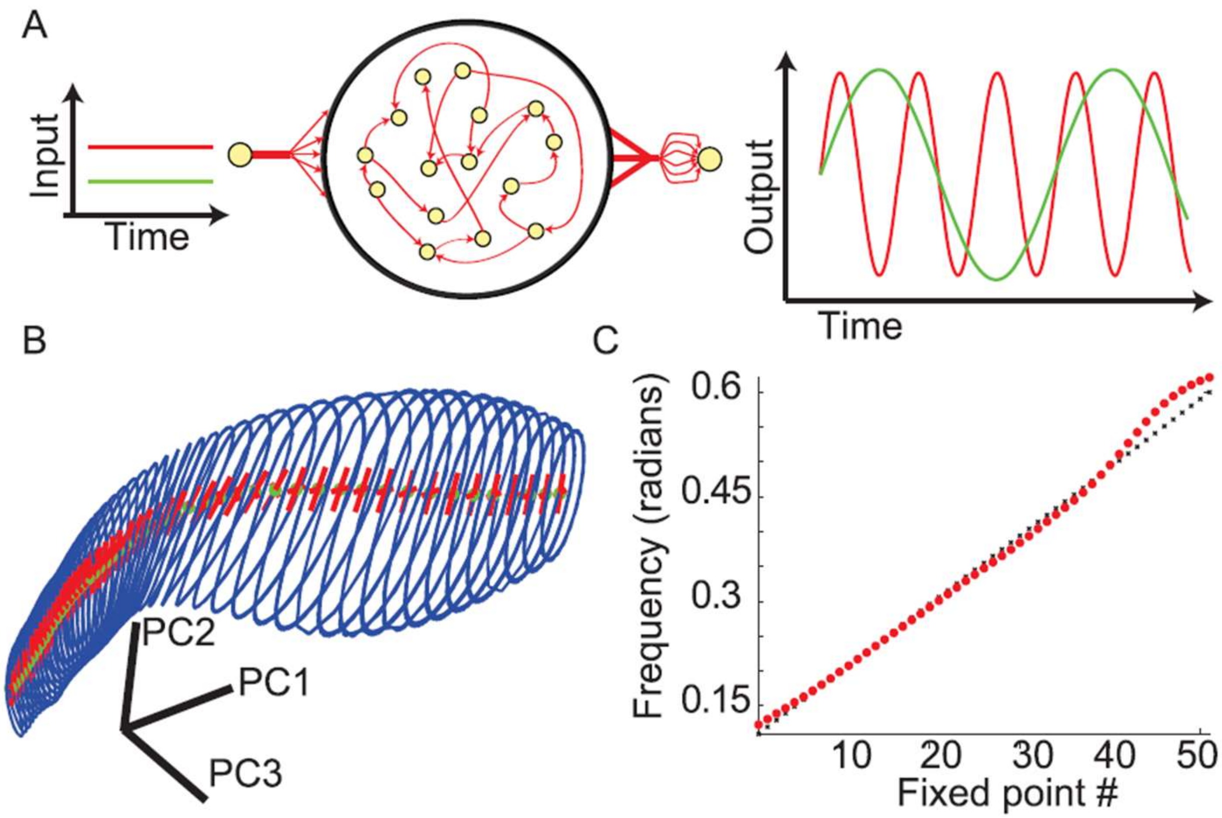


Figure 4 from D. Sussillo and O. Barak. Neural Computation **25** 626-649 (2013).

The reservoir (within reservoir computing) is a fixed random dynamical system.

It means that one nonlinear dynamical system can be used to emulate a different chaotic nonlinear dynamical system.

Theoretically it is based upon a "**fading memory function**".

Denote $u^{-\infty}(n) = (u(n), u(n-1), u(n-2), \ldots)$. $F(n)$ has fading memory if it depends less and less on $u(i)$ as $i \to \infty$.

Denote $F_m[u(n), u(n-1), \ldots, u(n-m)]$. F_m has fading memory if it converges to F as $m \to \infty$.

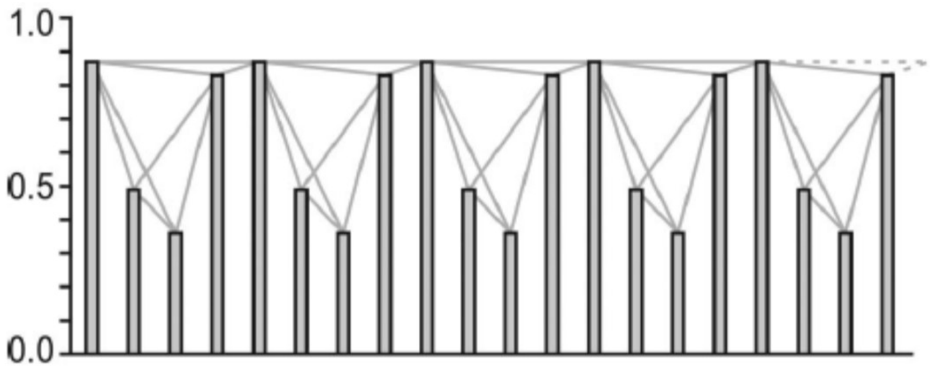
Many chaotic dynamical systems can be described with a recurrence relation that obeys this fading memory principle.

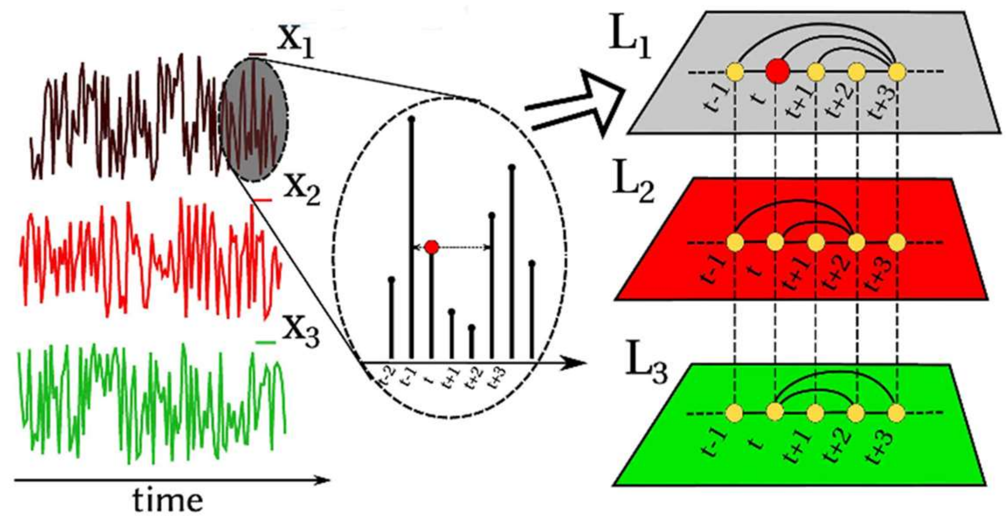
$$d(n) = D[d^{-\infty}(n-1)],$$

A reservoir computer has the same dynamical properties (Lyapunov exponent, stability regimes, etc.) as the system it is trained to replicate.

Ideas taken from P. Antonik, M. Gulina, J. Pauwels, and S. Massar. *Using a reservoir computer to learn chaotic attractors, with applications to chaos synchronization and cryptography.*

(*AN ASIDE…)*
How can we understand time series as graphs?

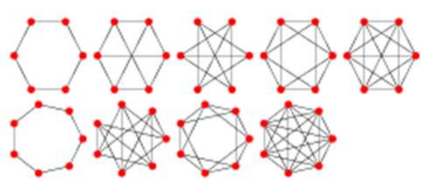




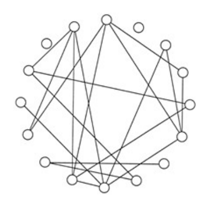
Here the rays between bars are the distances between vertices in a graph.

In this case, the network is undirected because the rays contain no information about time directionality.
But from a directional graph, you can learn something about conservation from time inversion (or lack of) in the graph.

Periodic time series → regular graphs



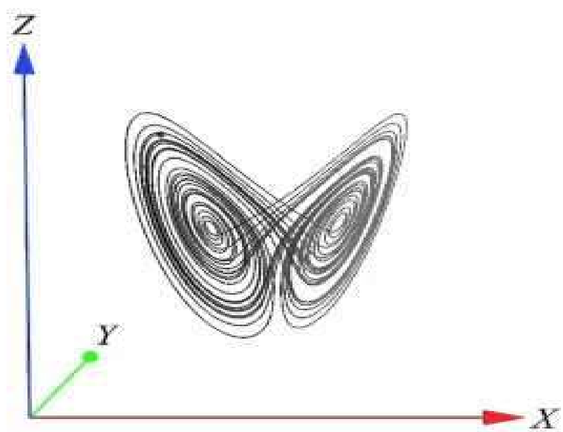
Stochastic series → random graphs



L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño. *From time series to complex networks: The visibility graph.* PNAS 105 (13) pp. 4972—4975 (2008).
L. Lacasa, V. Nicosia, and V. Latora. *Network structure of multivariate time series.* Sci. Rep. 5 (15508) (2015).

Empirical dynamical modeling is also **data-driven** and intended for **highly nonlinear** systems.

It builds up patterns from data, more formally called **attractor reconstruction.**



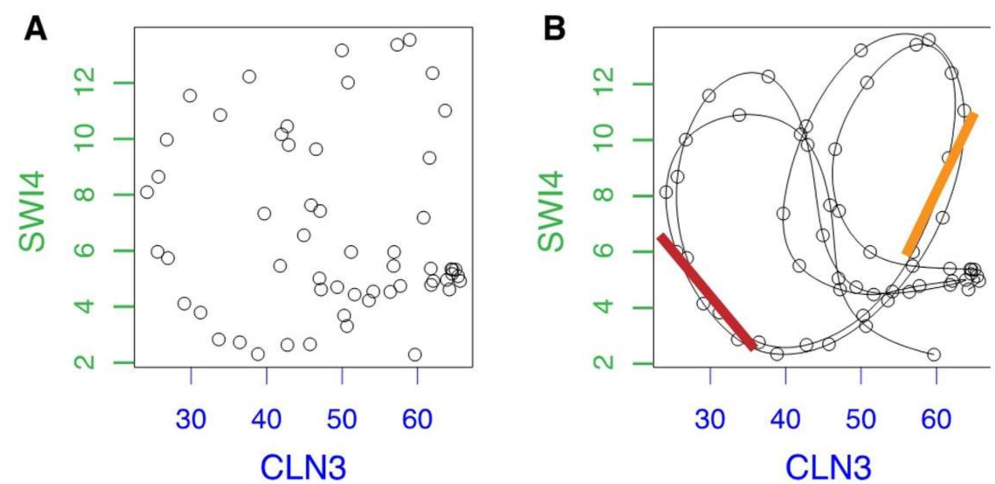


Image from http://deepeco.ucsd.edu/nonlinear-dynamics-research/edm/

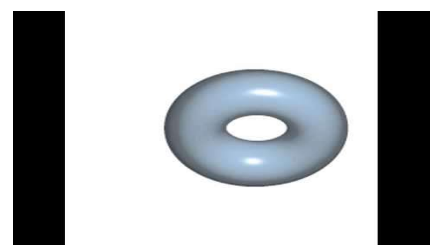
EDM is based upon Takens's theorem.

When can we reconstruct the manifold of a chaotic dynamical system using delayed time series or **delay embedding**?

If we have a time series $y(1), y(2), \ldots, y(N)$ from an *n*th order system represented by a *D*-dimensional attractor,

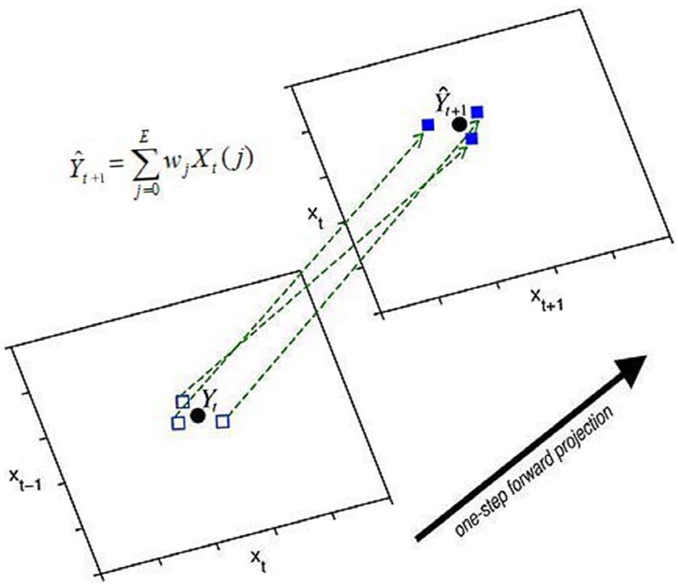
we can use the delay coordinates $y(k), y(k-\tau), \ldots, y(k-(d-1)\tau)$ with $d > 2n$ to reconstruct a non-unique shadow manifold.

The shadow manifold is **topologically invariant** from the original.



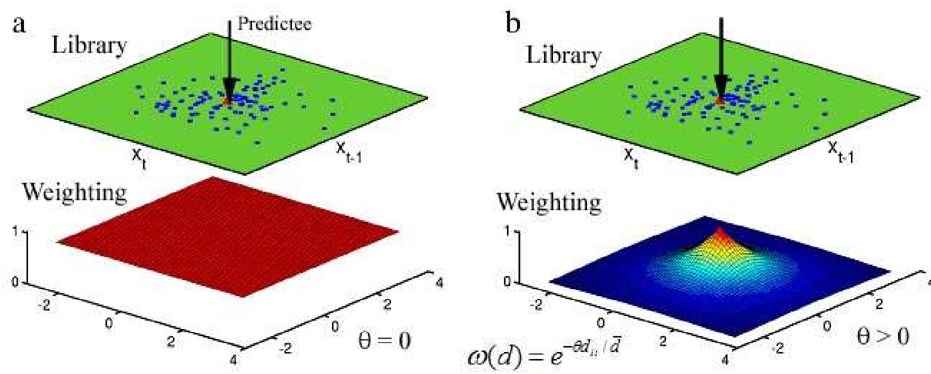
There are two main implementations of EDM.

Simplex projection



Take *n* points near the one you're forecasting on the attractor and forward step this "simplex" in time and take the average

S-map

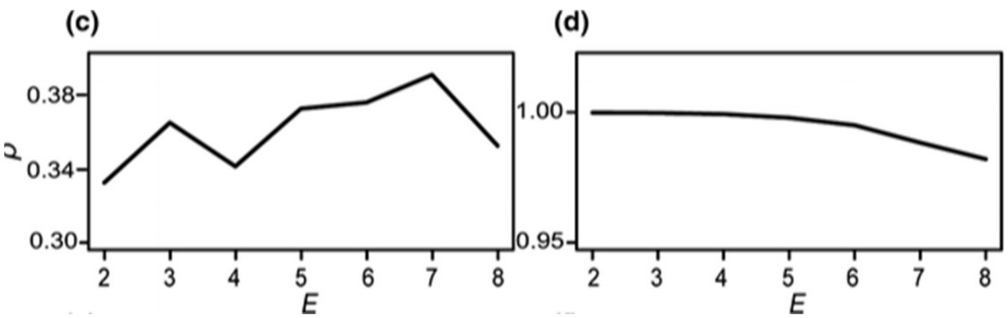


Take *n* points near the one you're forecasting, project these points, and fit a weighted curve (multidimensional spline) to their outcomes. The weighting includes a non-linearity parameter Θ

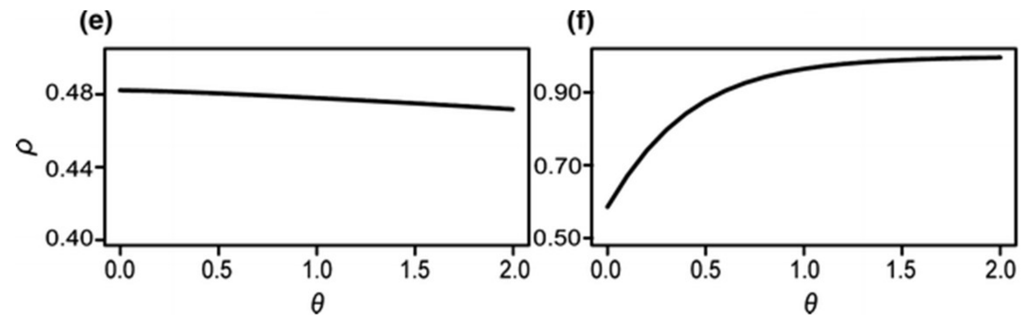
Images from http://deepeco.ucsd.edu/simplex/

EDM has a variety of uses:

(1) determine dimensionality of a system



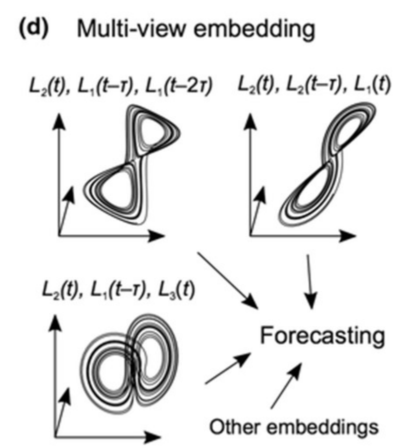
(2) quantify the degree of non-linearity in a system



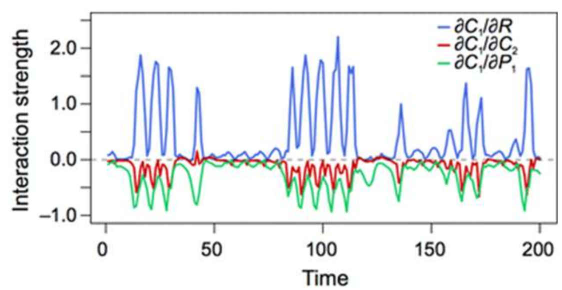
(3) determine causal variables

V_2 (t) to predict V_1 (t) ➔ V_1 caused V_2

(4) forecasting



(5) strength and sign of interaction



Chang et al. *Empirical dynamic modeling for beginners*. Ecological Research (2017) **32** No. 6 pp. 785-796.

COMPARISON…

RNN and reservoir computing

We use the network input now and the network output previously to calculate our hidden weights.

One nonlinear dynamical system can be used to emulate another.

EDM: simplex projection and S-maps

We use the delay coordinates $y(k), y(k-\tau), \ldots, y(k-(d-1)\tau)$ with $d > 2n$ to reconstruct a non-unique shadow manifold.

Forecasts for a subset of points on a manifold can be used to emulate another.

In summary,

Correlation is insufficient to determine causality in nonlinear dynamical systems.

Alternatives include

- Recurrent neural networks that use self loops and sequential data.

- Reservoir computers that use the fading memory principle and one dynamical system to emulate another.

- Empirical dynamical modeling that uses delay embedding to generate a shadow manifold from a subset of points.