EMBEDDING (1)

- It was suggested first in [1] that nothing really imposed that the vectors of the reconstructed attractor should be composed of equally spaced samples. Actually, what matters is the time span $(m\text{-}1)\tau$ covered by the vectors.
- But the approach proposed consists in selecting terms in a linear AR model and retaining those as vector components in the embedding.

EMBEDDING (2)

- In [2] this idea was somewhat amplified. Let us consider now embedded vectors

$$X_{n\text{-}1} = [x(n\text{-}i_1), x(n\text{-}i_2), \ldots, x(n\text{-}i_{\max})]^{\mathsf{T}}$$

with all $i_k < d$, d chosen beforehand, and one tries to predict sample $x(n)$ using some $G[X_{n\text{-}1}]$.

- Selection of the indices amounts to selecting a bit string a of length d.
- The idea is to select the best possible G and a using an MDL approach.

EMBEDDING (3)

- Recall that the coding length of data + model is:

$$L(x,\theta,a) = -\log P(x|\theta,a) + L(\theta) + L(a)$$

where θ is the model parameter vector.

- The log likelihood of the data $\{x(n)\}$, $n = 1, \ldots, N$, must include the log likelihood of the initial conditions $X^{(0)} = [x(d), x(2), \ldots, x(1)]^{\mathrm{T}}$:

$$-\log P(x|\theta,a) = -\log P(x|\theta,a,X^{(0)}) - \log P(X^{(0)})$$

EMBEDDING (4)

- If it is assumed that the errors and the initial conditions are independent and normally distributed, one gets:

$$L(x,\theta,a) = -\log P(x|\mathrm{N}(0,\sigma^2)) - \log P(X^{(0)}|\mathrm{N}(0,\sigma_X^{\,2})) + L(\theta) + L(a)$$

- Now, in addition to determining the best set of indices, one must also select the best model.

EMBEDDING (5)

- The idea proposed in [2] consists in using the deterministic prediction scheme as the model. Any imbedded vector can be used for prediction.
- This approach is simple to implement, and it is robust.
- It does not require any parameter, thus $L(\theta) = 0$.
- But it is only a pseudo-model, since prediction of $x(n)$ may imply X_k, with $k > n$.

EMBEDDING (6)

- After some manipulations, and suppression of constant or negligible terms, the MDL criterion is:

$$\mathrm{MDL}(a) \approx \frac{d}{2}\log\left[\frac{1}{d}\sum_{i=1}^{d}(x(i)-\bar{x})^2\right]$$

initial conditions

$$+\frac{N-d}{2}\log\left[\frac{1}{N-d}\sum_{i=d+1}^{N}e(i)^2\right]+d$$

model

errors

EMBEDDING (7)

- Examining all possible combinations of indices up to a maximum value *d* means estimating 2^d models. This is not feasible.
- What is proposed in [2] is a sequential selection procedure, i.e.:

- the first term is the one giving the smallest MDL and kept.
- All possible 2nd terms are tested and the best one kept if the MDL is smaller …

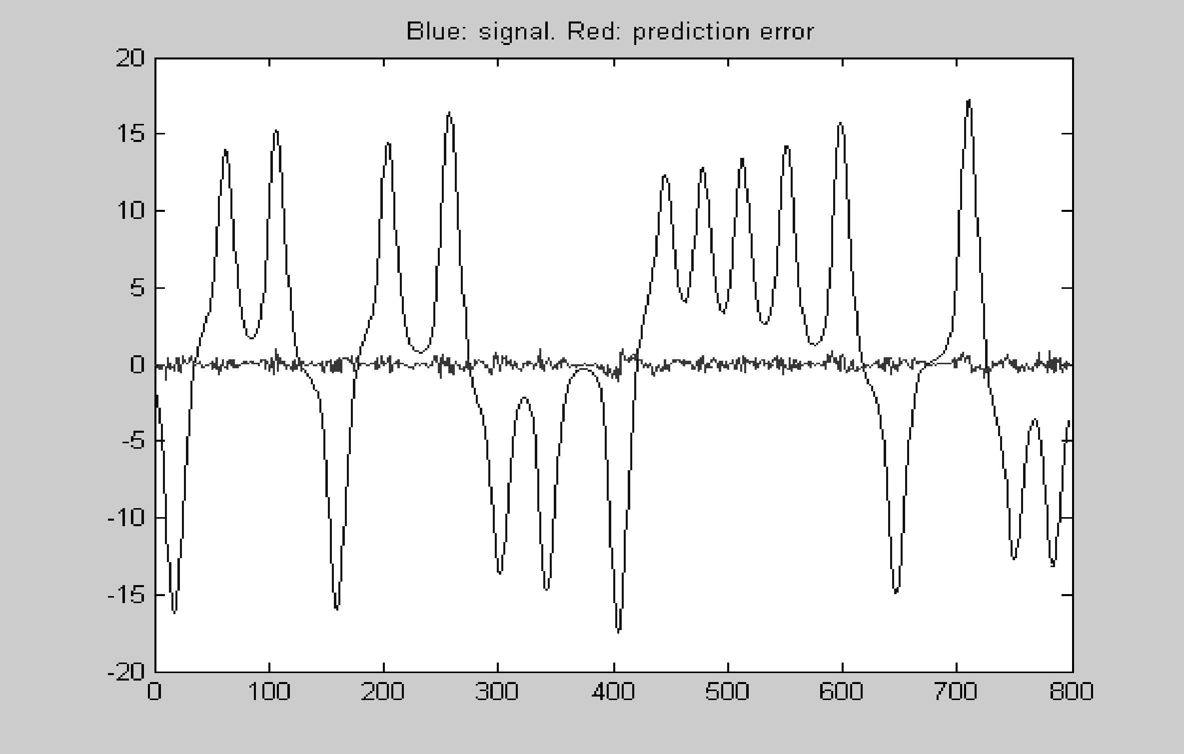
EMBEDDING (8)

- Unfortunately, this sequential selection does not work as well as for term selection in pseudo-linear models, which amounts to defining the best subspace to project to.
- Why not apply a genetic algorithm? Each possible embedding is coded as a bit string, with a ‘1’ if the component is selected, ‘0’ otherwise.

EMBEDDING (9)

- Lorenz system, *x* coordinate

Sequential and GA:
Same model
Indices 1, 2
(maximum 15)



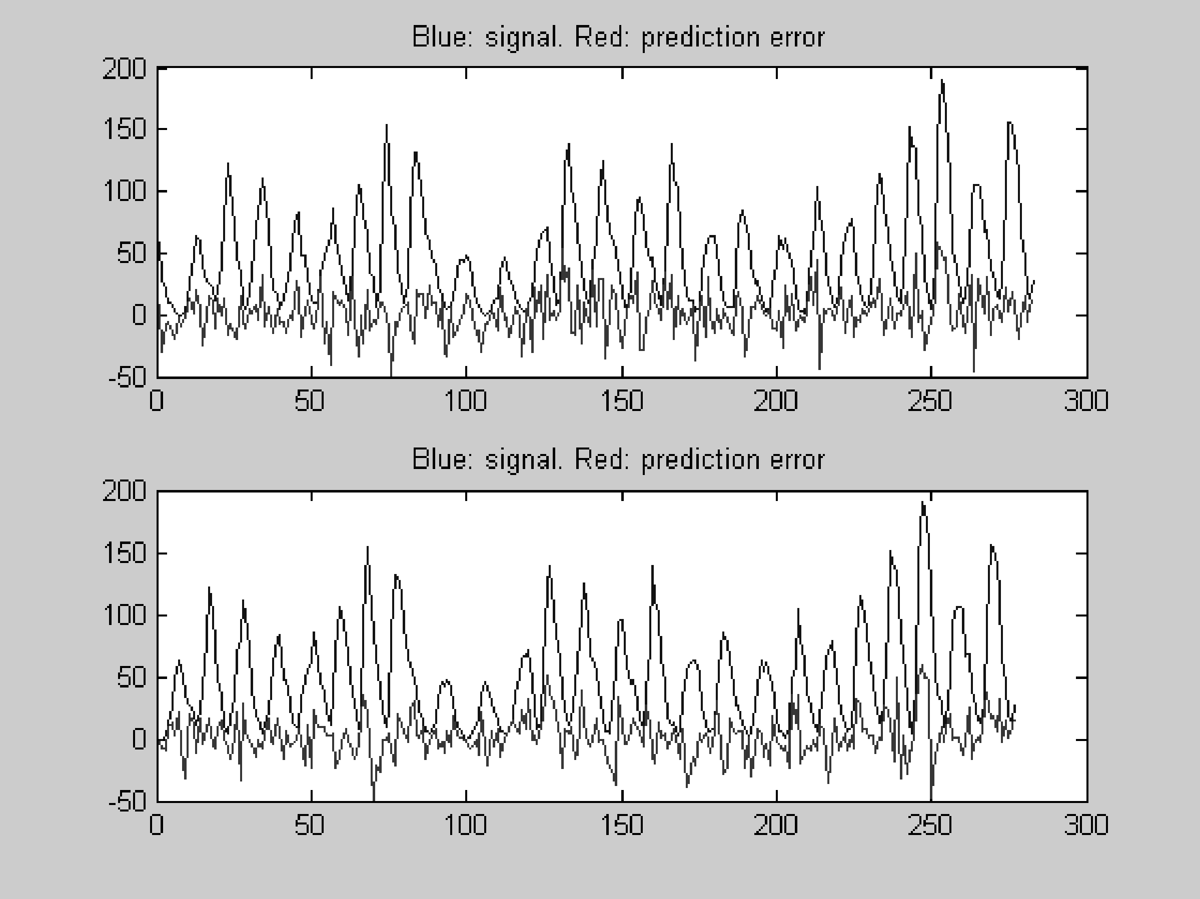
EMBEDDING (10)

- Sunspot time series

Sequential
Indices 1, 2, 5
MDL: 839.02

GA
Indices 1, 2, 3, 5, 11
MDL: 837.5

(maximum 15)



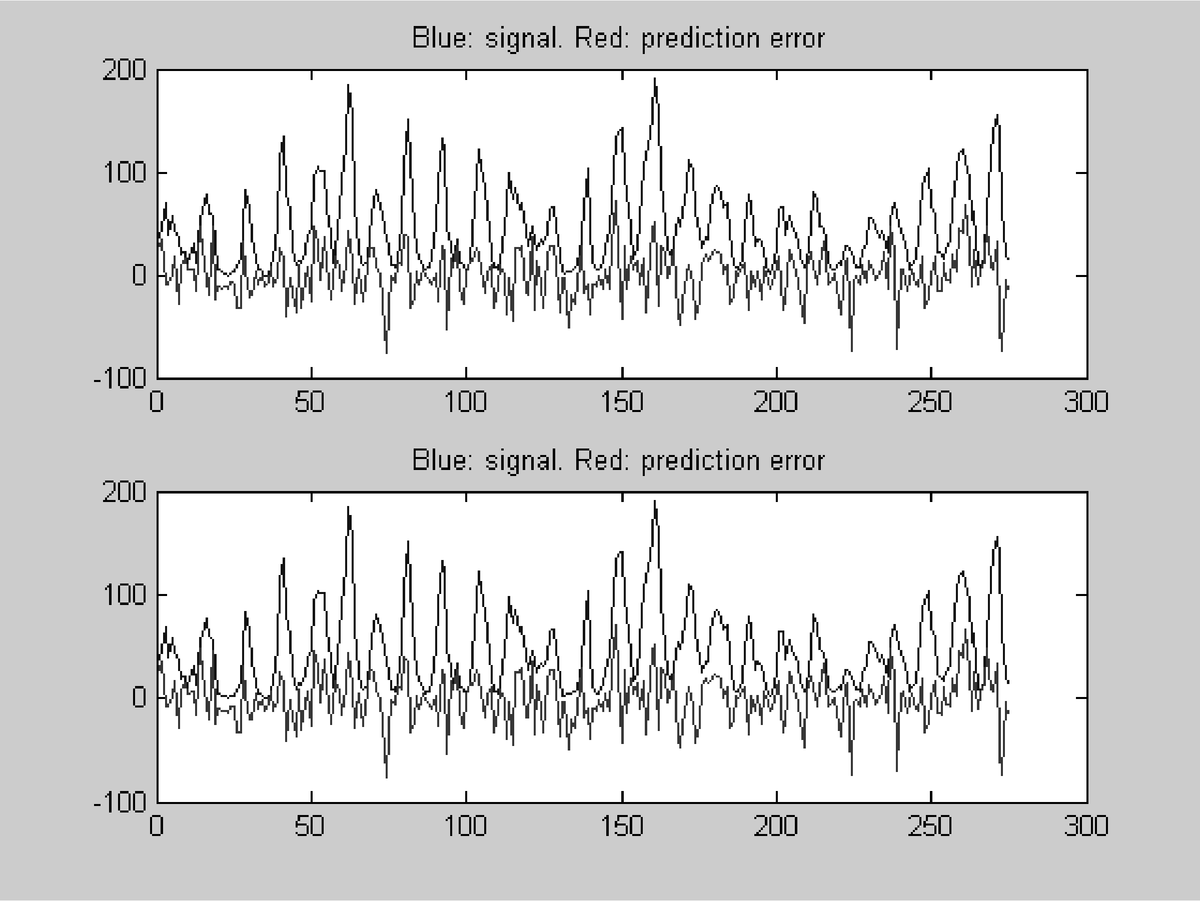
EMBEDDING (11)

- Surrogate on sunspot time series

Sequential
Indices 1, 3, 12, 13
MDL: 931.62

GA
Indices 1, 3, 12, 13
MDL: 931.62

(maximum 15)



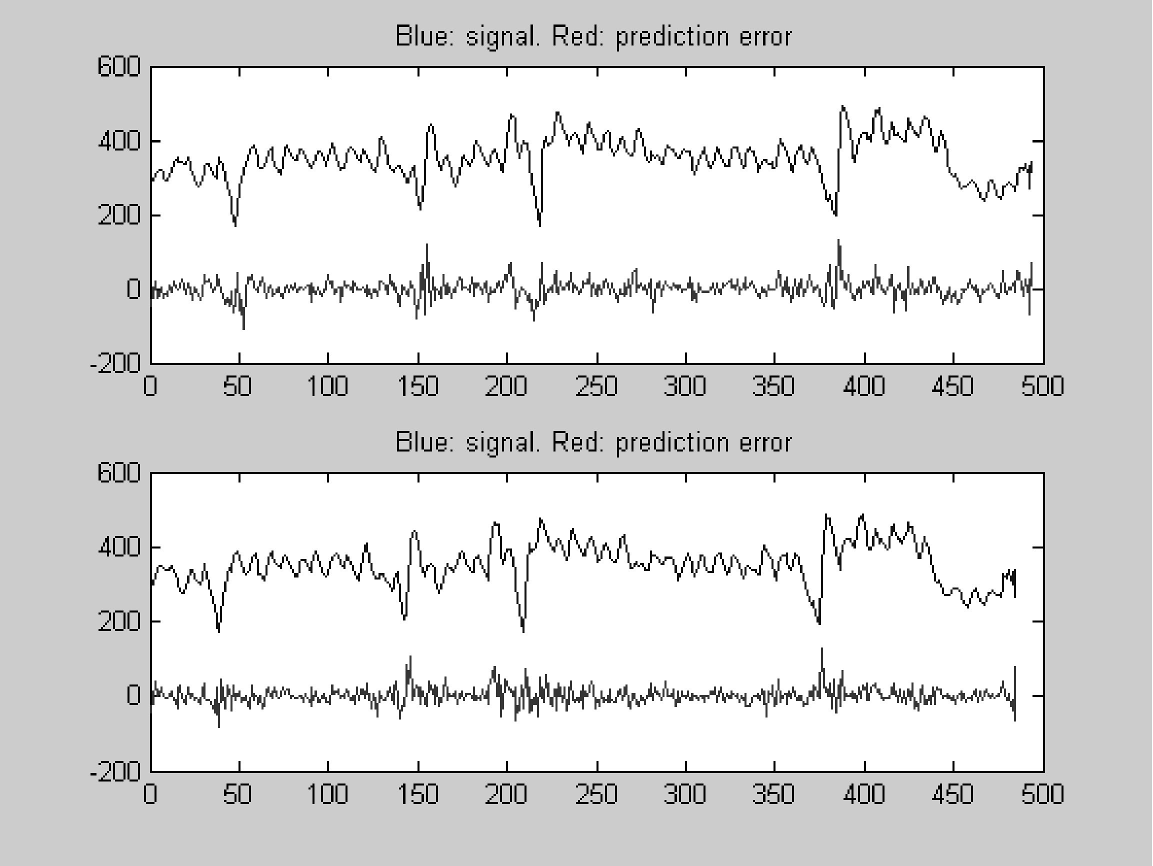
EMBEDDING (12)

- RR Intervals (between heartbeats), sampling 2 Hz.

Sequential
Indices 1, 3, 6
MDL: 1623.8

GA
Indices 1, 3, 12, 15
MDL: 1597.8

(maximum 15)



EMBEDDING (13)

- Increase in model size with noise (Henon map, 300 samples)

snr (dB)	terms
20	1, 2, 3
15	1, 2, 3
10	1, 2, 3, 8
5	1, 2, 3, 4, 7, 8, 12, 15
0	1, 2, 3, 5, 7, 9, 10, 11, 13, 14

INTERDEPENDENCE (1)

- A technique to assess the interdependence between two multivariate time series, related to deterministic prediction, has been proposed [3,4].
- Of course, those multivariate time series $\mathbf{X} = \{X_n\}$ and $\mathbf{Y} = \{Y_n\}$ may result from the embedding of two signals $\{x(k)\}$ and $\{y(k)\}$.
- The idea is to look for the possible existence of (well behaved) functions $G(.)$ and $F(.)$ such that

$$X_n = G(Y_n) \quad \text{and} \quad Y_n = F(X_n)$$

INTERDEPENDENCE (2)

- It is to be noted that interdependence does not imply causation. The relationship between the two series may be due to the influence of a third one on both of them.
- The principle of the test is simple and robust: if indeed $G(.)$ and $F(.)$ exist and are reasonably continuous, then if Y_n is close to Y_k, then $X_n = G(Y_n)$ should be close to $X_k = G(Y_k)$.
- Closeness is quantified with respect to a deterministic prediction of X_n using closest neighbors in $\{X_k\}$.

INTERDEPENDENCE (3)

- Let $r_{n,j}$ and $s_{n,j}$, $j = 1, \ldots, k$, denote the indices of the k nearest neighbors of X_n and Y_n respectively.
- The mean squared distance between X_n and its closest neighbors is:

$$R_n^{(k)}(\mathbf{X}) = \frac{1}{k} \sum_{j=1}^{k} \| X_n - X_{r_{n,j}} \|^2$$

This quantity should be small for close enough neighbors.

INTERDEPENDENCE (4)

- The *conditional* mean squared distance with respect to $\{Y_k\}$ is:

$$R_n^{(k)}(\mathbf{X}|\mathbf{Y}) = \frac{1}{k}\sum_{j=1}^{k} \| X_n - X_{s_{n,j}} \|^2$$

- One can define the same quantities for Y_n :

$$R_n^{(k)}(\mathbf{Y}) = \frac{1}{k}\sum_{j=1}^{k} \| Y_n - Y_{s_{n,j}} \|^2 \qquad R_n^{(k)}(\mathbf{Y}|\mathbf{X}) = \frac{1}{k}\sum_{j=1}^{k} \| Y_n - Y_{r_{n,j}} \|^2$$

INTERDEPENDENCE (5)

- If the point set $\{Y_k\}$ has an average squared radius $R(\mathbf{X})$:

$$R_n^{(k)}(\mathbf{X}) \big/ R(\mathbf{X}) \approx (k/N)^{2/m} << 1 \text{ for } k << N$$

with m the dimension and N the number of vectors.

- If **X** and **Y** are indeed interdependent, then the same relationship should hold for the conditional mean square distance.

INTERDEPENDENCE (6)

- A measure of local interdependence is:

$$S_n^{(k)}(\mathbf{X}|\mathbf{Y}) = \frac{R_n^{(k)}(\mathbf{X})}{R_n^{(k)}(\mathbf{X}|\mathbf{Y})} \qquad 0 < S_n^{(k)}(\mathbf{X}) \leq 1$$

- And the global measure is:

$$S^{(k)}(\mathbf{X}|\mathbf{Y}) = \frac{1}{N}\sum_{n=1}^{N} S_n^{(k)}(\mathbf{X}|\mathbf{Y}) \qquad 0 < S^{(k)}(\mathbf{X}|\mathbf{Y}) \leq 1$$

- This global measure should be close to 0 when there is no interdependence, and close to 1 for strong interdependence

INTERDEPENDENCE (7)

- To assess the significance of $S^{(k)}(\mathbf{X}|\mathbf{Y})$, one can generate several surrogate data $\mathbf{Y}^{(u)}$ for $\mathbf{Y}$, compute $S^{(k)}(\mathbf{X}|\mathbf{Y}^{(u)})$, and perform a rank test.
- If there is indeed a significant interdependence, it is possible to test if it is linked to phase only by generating bivariate surrogated data $\{\mathbf{X}^{(b)},\mathbf{Y}^{(b)}\}$, compute $S^{(k)}(\mathbf{X}^{(b)}|\mathbf{Y}^{(b)})$, and perform a rank test.
- Those bivariate surrogate data are generated by applying the same phase randomization to both signals.

RECURRENCE PLOT (1)

- Recurrence plots (RP) were initially introduced to display the recurrence of patterns and possible non stationarities in imbedded time series [5,6].
- Apart from producing nice pictures, several parameters can be extracted from RP to characterize dynamical processes.
- RP have also been generalized to the analysis of the interdependence between two time series.

RECURRENCE PLOT (2)

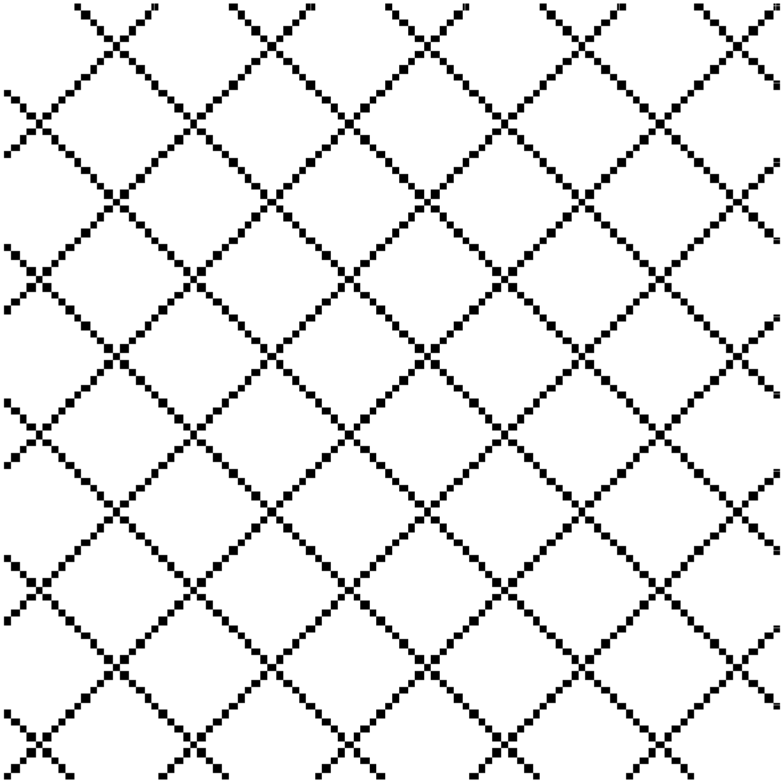
- The principle of RP is quite simple: given a multivariate time series $\{X_n\}$ $n = 1, \ldots, N$, possibly obtained through the embedding of a signal, an RP is an $N \times N$ array in which a dot is placed at location (i,j) if $\|X_i - X_j\| < \delta$, δ a predetermined small number.
- Typically, δ is a fraction of the sum of standard deviations of the vector components (sstd).
- Of course, RPs are symmetrical with respect to the main diagonal.

RECURRENCE PLOT (3)

- Periodicities (recurrences) are expressed by diagonal lines in RPs. RP-based analysis is mostly based on the characterization of these lines.
- Of course, an RP is dependent upon the threshold δ, but it is also quite dependent upon vector dimension *m*.
- What matters is to conserve these values through the various experiments performed.

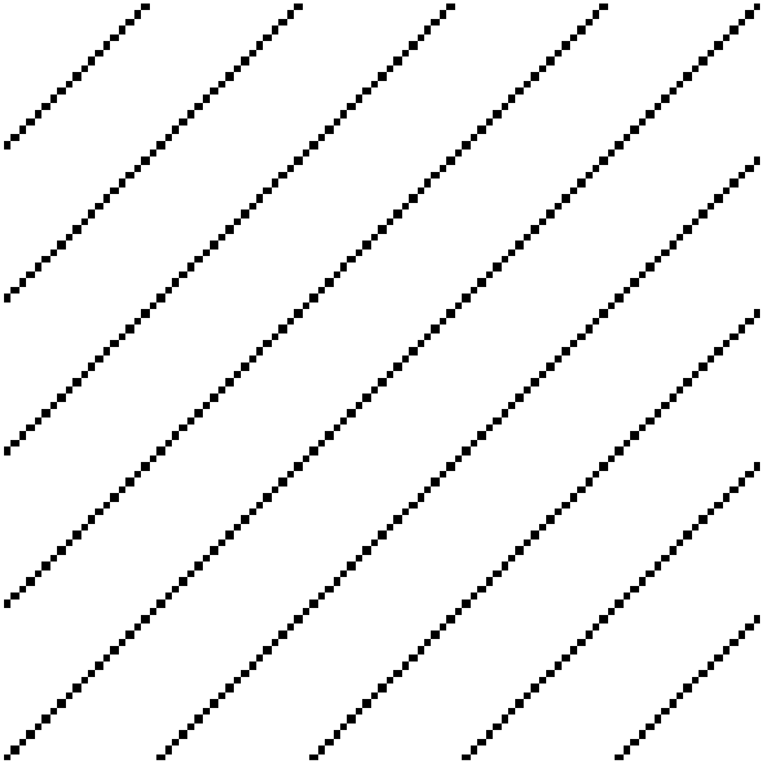
RECURRENCE PLOT (4)

- Example: sinusoid with period 20, *m* = 1, δ = 0.05 sstd.



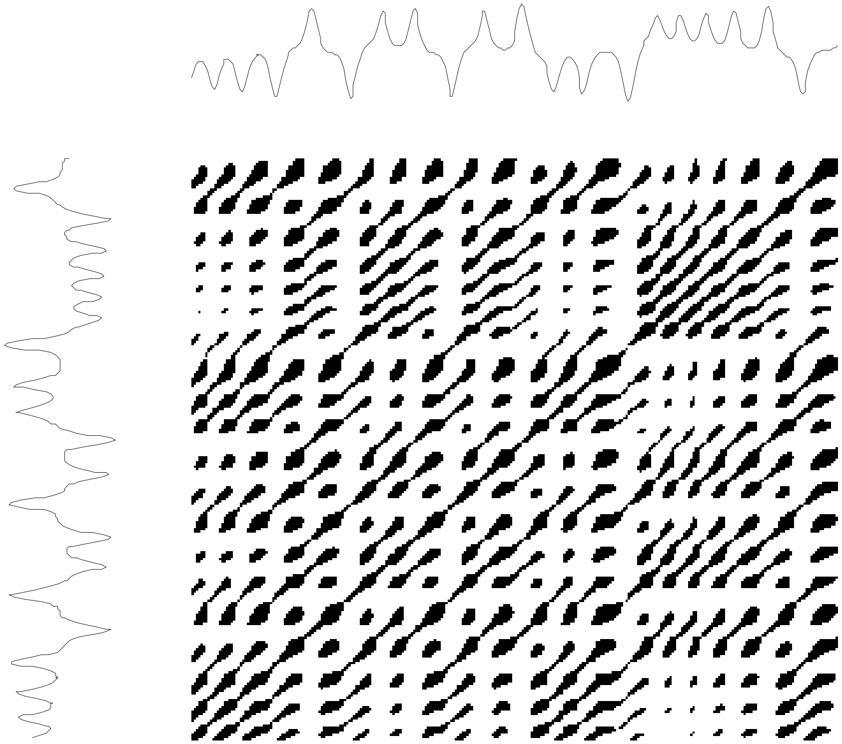
RECURRENCE PLOT (5)

- Example: sinusoid with period 20, *m* = 2, δ = 0.05 sstd.



RECURRENCE PLOT (6)

- Example: Lorenz, [*x y z*], δ = 0.05 sstd.



RECURRENCE PLOT (7)

Parameters extracted:

- REC proportion of recurrence (black) points. Higher for periodic dynamics.
- DET proportion of recurrence points in diagonal lines of length at least 2. Higher for deterministic dynamics.
- ENT entropy of line lengths on the diagonal. Higher for more complex dynamics

RECURRENCE PLOT (8)

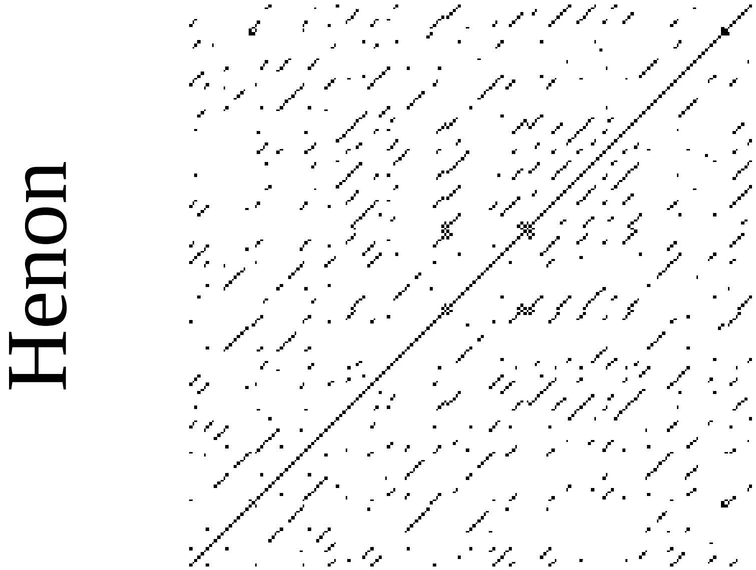
- DIV inverse of the length of the longest diagonal line. Higher for larger maximum Lyapunov exponent.
- TREND slope *a* of the regression line:

$$R_j = b + ad_j + \varepsilon_j$$

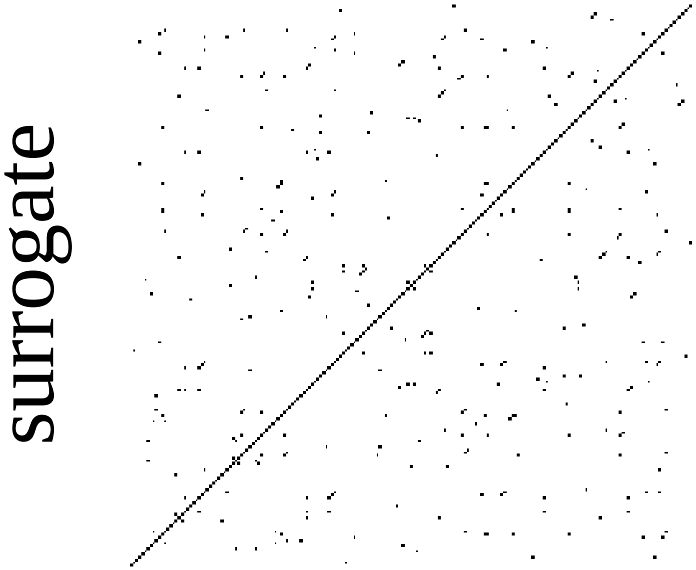
with R_j the proportion of recurrence points on diagonal at distance d_j from the main diagonal. Quantifies the decrease in recurrence with respect to time difference.

RECURRENCE PLOT (9)

- Clearly, significance of the values of these parameters in terms of determinism or nonlinearity must be assessed using surrogate data.



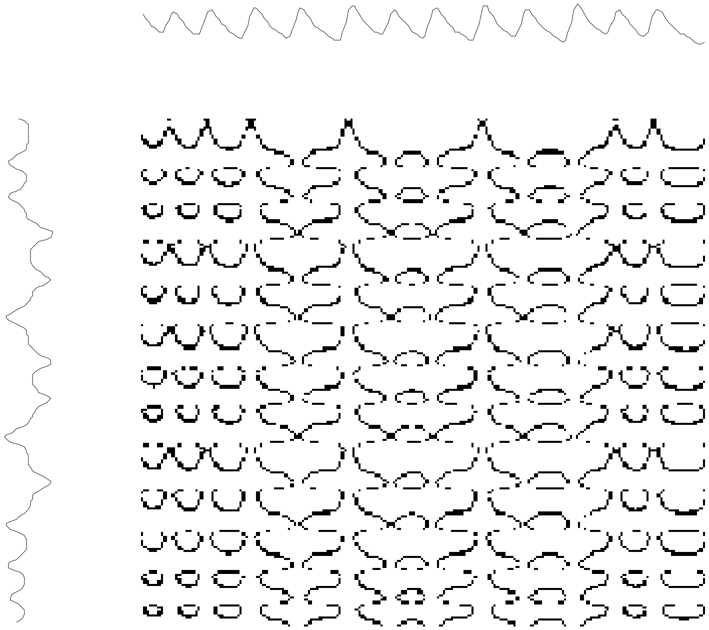
REC=0.03 DET=0.86 DIV=0.09 trend=-1.5e-5



REC=0.008 DET=0.24 DIV=0.33 trend=-1.7e-5

RECURRENCE PLOT (10)

- It is also possible to generate RP with two time series (same range, same dimension).
- Example: x and *y* of Lorenz



REFERENCES

1. K. Judd and A. Mees ˝Embedding as a modeling problem,˝ *Physica D*, vol. 120, 1998, pp. 273-286.
2. M. Small and C.K. Tse, ˝Optimal embedding parameters,˝ *Physica D*, vol. 194, 2004, pp. 283-296.
3. J. Arnhold, P. Grassberger, K. Lehnertz, and C.E. Elger, ˝A robust method for detecting interdependencies: Application to intracranially recorded EEG,˝ *Physica D*, vol. 134, 1999, pp. 419-430.
4. E. Pereda, R. Rial, A. Gamundi, and J. Gonzalez, ˝Assessement of changing interdependencies between human electroencephalograms using nonlinear methods,˝ *Physica D*, vol. 148, 2001, pp. 147-158.
5. J. Giao and H. Cai, ˝On the structures and quantification of recurrence plots,˝ *Phys. Lett. A*, vol. 270, 2000, pp. 75-87.
6. J. Belaire-Franch, D. Contreras, and L. Tordera-Lledo, ˝Assessing nonlinear structures in real exchange rates using recurrence plot strategies,˝ *Physica D*, vol. 171, 2002, pp. 249-264.