STATE SPACE RECONSTRUCTION FROM MULTIPLE TIME SERIES

I. VLACHOS* and D. KUGIUMTZIS

Department of Mathematical, Physical and Computation Sciences,
Faculty of Engineering Aristotle University of Thessaloniki,
Thessaloniki, 54124, Greece
**E-mail: ivlaxos@gen.auth.gr*
web.auth.gr/gen

In the non-linear analysis of scalar time series the common practice is to reconstruct the state space using time-delay embedding. When there are more than one observed quantities, one can reconstruct the state space using a time-delay embedding scheme specifying embedding parameters for each quantity. In this study we investigate the state space reconstruction from multiple time series derived from continuous systems and propose a method for building the embedding vector progressively using information measure criteria. The proposed method is compared to other methods with simulations on known chaotic systems, such as individual and coupled Lorenz and Rössler systems. Our analysis showed that multivariate attractor reconstruction preserves better the dynamics of a system and our proposed method gives a parsimonious alternative to the simple extension of the univariate case.

Keywords: Non-linear analysis; Multivariate time series; State space reconstruction.

1. Introduction

Since its publication, Takens' Embedding Theorem[1,2] is reputed as the most significant tool in non-linear time series analysis and has been used in many different settings ranging from system characterization and approximation of invariant quantities to prediction and noise-filtering.[3] The Embedding Theorem implies that although the true dynamics of a system may not be known, equivalent dynamics can be obtained using time delays of a single time series, seen as the one-dimensional projection of the system trajectory.

Most applications of the Embedding Theorem deal with univariate time series, but often measurements of more than one quantities related to the same dynamical system are available. One of the first uses of multivariate embedding was in the context of spatially distributed systems where embedding vectors were formed from simultaneous measurements of a variable at different locations.[4,5] Multivariate embedding was used among others for noise reduction[6] and surrogate data generation.[7] Another particular implementation was the prediction of a time series using local models on a state space reconstructed from a different time series of the same system.[8] In a later work, embedding was extended to all of the observed time series,[9] and recently non-uniform multivariate embedding using different delays was briefly addressed in Ref. 10,11. All of these works are basically extensions of methods used in the scalar case with slight modifications without taking into account particular characteristics of multivariate embedding, such as possible high dependence of individual components obtained from the different time series.

In this work, we contemplate on the particularities of multivariate embedding and their effect on the state space reconstruction and suggest a new embedding scheme to treat them. We study only time series generated from continuous dynamical systems where complications mostly occur. Based on Fraser's idea of using information criteria to select the embedding parameters in the univariate case[12,13] we propose a method for progressive building of the embedding vectors allowing for different variables and delays. We compare our proposed method with the univariate reconstruction on each single time series and the direct extension to all the time series with respect to the distortion induced on the original attractor under the embedding processes. Simulations on known chaotic systems are used to evaluate the embedding techniques.

In Sec. 2, the embedding for univariate time series is briefly discussed. In Sec. 3, the discussion is extended to multivariate time series and the proposed embedding technique is presented, and in Sec. 4, the tools for comparing the reconstructions are given. In Sec. 5, the results of simulations are presented and, finally, in Sec. 6, the results are discussed and conclusions are given.

2. Univariate Embedding

A continuous dynamical system generates a trajectory in a D-dimensional manifold Γ and its evolution can be described by D ordinary differential equations

$$\dot{\mathbf{y}} = f(\mathbf{y}),$$

where $\mathbf{y} \in \Gamma$ is the state vector at time t.

The observed scalar time series $\{x_n\}_{n=1}^N$ of length N is the projection of the state vector of the system at discrete times $\{\mathbf{y}_n\}_{n=1}^N$ given by a measurement function $h : \Gamma \mapsto \mathbb{R}$ as $x_n = h(\mathbf{y}_n)$. Besides the apparent loss of information of the system dynamics by the projection, the system dynamics may be recovered through suitable state space reconstruction from the single time series.

Assume that the dynamical system "lives" on an attractor $A \subset \Gamma$. According to Taken's embedding theorem a trajectory formed by points $\mathbf{x}_n$ of the delay map

$$\mathbf{x}_n = (h(\mathbf{y}_n), h(\mathbf{y}_{n-\tau}), ..., h(\mathbf{y}_{n-(m-1)\tau})) = (x_n, x_{n-\tau}, ..., x_{n-(m-1)\tau}), \quad (1)$$

under certain genericity assumptions, is an one-to-one image of the original trajectory of $\mathbf{y}_n$ provided that m is large enough.

The reconstructed attractor $\tilde{A}$ obtained by the time-delay vectors is topologically equivalent to A as long as $m \geq 2d + 1$ where d is the box-counting dimension of A, and invariant quantities of the original system (dimensions, Lyapunov exponents, etc) can be estimated from the reconstructed dynamics.

The embedding process is visualized in the following diagram where $\mathbf{F}$ is the original map at discrete steps, e is the embedding procedure creating the delay vectors from the time series and $\mathbf{G}$ is the reconstructed dynamical system on $\tilde{A}$.

$$\begin{tikzcd}
\mathbf{y}_n \in A \subset \Gamma \arrow[r, mapsto, "\mathbf{F}"] \arrow[d, "h"'] & \mathbf{y}_{n+1} \in A \subset \Gamma \arrow[d, "h"'] \\
x_n \in \mathbb{R} \arrow[d, "e"'] & x_{n+1} \in \mathbb{R} \arrow[d, "e"'] \\
\mathbf{x}_n \in \tilde{A} \subset \mathbb{R}^m \arrow[r, mapsto, "\mathbf{G}"] & \mathbf{x}_{n+1} \in \tilde{A} \subset \mathbb{R}^m
\end{tikzcd}$$

The two parameters of the delay embedding in (1) are the embedding dimension m, i.e. the number of components in $\mathbf{x}_n$ and the delay time τ. Among the approaches for the selection of m the most popular is the *false nearest neighbors* (FNN)[14] while τ is usually chosen as the one that gives the first minimum of the time-delayed mutual information (MI).[15]

3. Multivariate Embedding

Given that there are p time series $\{x_{i,n}\}_{n=1}^N$, $i = 1, \ldots, p$, generated through p measurement functions $h_i(\mathbf{y}_n)$ from the same system, the simplest extension of the reconstructed state space vector in (1) is of the form

$$\mathbf{x}_n = (x_{1,n}, x_{1,n-\tau}, \ldots, x_{1,n-(m-1)\tau}, x_{2,n}, \ldots, x_{p,n-(m-1)\tau}). \qquad (2)$$

The corresponding diagram for the multivariate embedding process is given below, where the dimension of the reconstructed space is M

$$
\begin{tikzcd}
 & \mathbf{y}_n \in A \subset \Gamma \arrow[dl, "h_1"'] \arrow[d, "h_2"', "\cdots"] \arrow[dr, "h_p"] \arrow[rrrr, mapsto, "\mathbf{F}"] & & & & \mathbf{y}_{n+1} \in A \subset \Gamma \arrow[dl, "h_1"'] \arrow[d, "h_2"', "\cdots"] \arrow[dr, "h_p"] & \\
x_{1,n} \arrow[dr, "e"'] & x_{2,n}\cdots \arrow[d, "e"', "\cdots"] & x_{p,n} \arrow[dl, "e"] & & x_{1,n+1} \arrow[dr, "e"'] & x_{2,n+1}\cdots \arrow[d, "e"', "\cdots"] & x_{p,n+1} \arrow[dl, "e"] \\
 & \mathbf{x}_n \in \tilde{A} \subset \mathbb{R}^M \arrow[rrrr, mapsto, "\mathbf{G}"] & & & & \mathbf{x}_{n+1} \in \tilde{A} \subset \mathbb{R}^M &
\end{tikzcd}
$$

A question that arises immediately is if all the information contained in the reconstructed vectors is useful. Some time series may be contemporary and/or lagged dependent on others and then the embedding vectors carry redundant information. The inclusion of such components just complicates the state space reconstruction and this may have implications in the analysis, e.g. degradation of the estimates of measures evaluated on the reconstructed attractor.

Judd and Mees[16] introduced the non-uniform embedding scheme for univariate time series using unequal lags, so that $\mathbf{x}_n = (x_{n-l_1}, x_{n-l_2}, \ldots, x_{n-l_m})$. The use of unequal lags may reduce better the redundant information than the embedding with fixed lag. Extending this to multivariate embedding we get

$$\mathbf{x}_n = (x_{1,n-l_{11}}, x_{1,n-l_{12}}, \ldots, x_{1,n-l_{1m_1}}, x_{2,n-l_{21}}, \ldots, x_{p,n-l_{pm_p}}).$$

Inspection of all possible combinations of components for determination of the optimum embedding (with any criterion we may choose) can be computationally intractable when the embedding dimension and number of time series are large. For a moderate example when $p = 3$, $m = 3$ and the maximum lag for all time series is 10 we have 4050 cases, while when we increase m to 4 we have 27405 cases. It is evident that for most practical purposes we need a progressive method to build the embedding vectors.

Assume that we have somehow chosen the maximum lags L_i for each time series $i = 1, \ldots, p$. Then we have a set of $L_1 \times L_2 \times \ldots \times L_p$ candidate components to choose from. Using ideas developed for the state space reconstruction from univariate time series, the reconstructed vector must satisfy two properties, first its components must be least dependent to each other and secondly it must be able to explain best the future states of the system for one, two or several steps ahead. In the following we refer to only a single, immediate future state.

Starting from an empty vector $\mathbf{b}_0$, at each step j that we have already selected the $j-1$ components $\mathbf{b}_{j-1} = (x_1^\star, x_2^\star, \ldots, x_{j-1}^\star)$ we add the lag

$x_j^\star \in \{x_{1,n}, x_{1,n-1}, \dots, x_{1,n-L_1}, x_{2,n}, \dots, x_{p,n-L_p}\} \setminus \{x_1^\star, \dots, x_{j-1}^\star\}$ that fulfills the following maximization criterion

$$\max_{x_j^\star} \Big\{ \frac{1}{p} \sum_{i=1}^{p} \mathrm{I}\big(x_{i,n+1}; \big(x_j^\star, \mathbf{b}_{j-1}\big)\big) - \mathrm{I}\big(x_j^\star; \mathbf{b}_{j-1}\big) \Big\}. \tag{3}$$

The first mutual information term in the sum gets large when the new embedding vector (with the addition of $x_j^\star$) explains best the dynamics of the system, while the second term is small when the candidate component is independent to the already included components, so maximization of (3) leads to embedding vectors that satisfy the two desired properties. Note that in essence the components of the delay vectors can never be fully independent since the two mutual information terms are related and the maximization gives actually the best tradeoff between independency of the components and explaining the future of the system.

The mutual information in (3) regards vector variables and thus its estimation requires more data than for the standard case of scalar variables. Data demand increases dramatically with the dimension, at the level of millions of data points for large dimension. Thus we approximate multidimensional mutual information in (3) with average pair-wise mutual information as

$$\mathrm{I}\big(x_j^\star; \mathbf{b}_{j-1}\big) \approx \frac{1}{j-1} \sum_{l=1}^{j-1} \mathrm{I}\big(x_j^\star; x_l^\star\big),$$
$$\mathrm{I}\big(x_{i,n+1}; \big(x_j^\star, \mathbf{b}_{j-1}\big)\big) \approx \frac{1}{j-1} \sum_{l=1}^{j-1} \mathrm{I}\big(x_{i,n+1}; \big(x_j^\star, x_l^\star\big)\big).$$

We use the standard mutual information estimate of equiprobable binning, which was found to perform satisfactorily among other estimates.[17]

For the determination of L_i we use the *Minimum Description Length*[18] selection criterion on each i time series separately, as was used by Judd and Mees in Ref.16. As a stopping criterion for the scheme of progressive vector building we use the idea of FNN. The expansion of $\mathbf{b}_{j-1}$ terminates when the percentage of false neighbors going from $\mathbf{b}_{j-1}$ to $\mathbf{b}_j$ with the selected component added at step j is smaller than a properly chosen threshold. The final reconstructed vector is then $\mathbf{b}_{j-1}$.

4. Comparing reconstructions

The reconstructed attractor is a distorted "image" of the original one, and the distortion may be with respect to its shape or its dynamics. The better the reconstruction, the less the distortion of the attractor. Thus we base the comparison of different reconstructions on two distortion measures, one for the shape and one for the dynamics.

Supposing that A is the original attractor and $\tilde{A}$ the reconstructed attractor, we measure the shape distortion with *Mutual False Nearest Neighbors*.[19] Let $\mathbf{y}_n$ be a point on A with $\mathbf{y}_{n_k}$ its nearest neighbor and respectively $\mathbf{x}_n$ the corresponding point on $\tilde{A}$ with nearest neighbor $\mathbf{x}_{n_r}$. The shape distortion is given by

$$Sd = \left\langle \frac{\| \mathbf{y}_n - \mathbf{y}_{n_k} \|}{\| \mathbf{x}_n - \mathbf{x}_{n_k} \|} \frac{\| \mathbf{x}_n - \mathbf{x}_{n_r} \|}{\| \mathbf{y}_n - \mathbf{y}_{n_r} \|} \right\rangle,$$

where $< . >$ is the average over all points and $\| . \|$ is the Euclidian distance.

To measure the distortion of the dynamics we use the *Normalized Root Mean Squared Error* of the one-step ahead nearest neighbor prediction on the original attractor. If $\mathbf{x}_n$ is a point on the reconstructed attractor with nearest neighbor $\mathbf{x}_{n_r}$, then the dynamics distortion is given by

$$Dd = \sqrt{\frac{\sum_n \| \mathbf{y}_{n+1} - \mathbf{y}_{n_r+1} \|^2}{\sum_n \| \mathbf{y}_{n+1} - \bar{\mathbf{y}} \|^2}},$$

where the summation is over all points and $\bar{\mathbf{y}}$ is the center of mass of the points in A. The larger Sd is and the smaller Dd is, the less distortion there is on the shape and on the dynamics of the attractor respectively. Sd is a stricter criterion. The dynamics may be preserved though the reconstructed attractor is a stretched or squeezed (or both) version of the original attractor. On the other hand, since the shape of the attractor is formed by the dynamics, the preservation of the shape guarantees the preservation of the dynamics.

An example for the dependence of the measures on the embedding is given in Fig. 1 for the reconstruction of the Lorenz attractor with the x and z variables using non-uniform multivariate embedding with $L_i = 10, i = 1, 2$ and for embedding dimensions 3 and 4. All possible combinations where taken and the measure values where sorted for clarity. We notice that there are more than one "optimal" (with respect to the dynamics distortion) embeddings for a given embedding dimension m that unfolds the attractor and there are even more "good" embeddings for larger m. For $m = 3$ we have approximately 1200 possible embeddings and 40 with Dd less than 0.03 that become 5000 and 400, respectively when m increases to 4. This fact exhibits the problem of identification that appears when dealing with multiple time series.

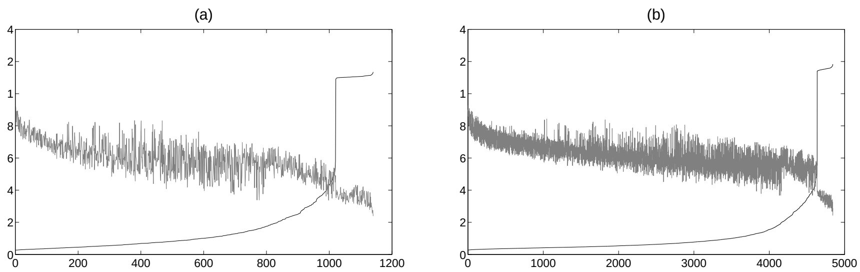

Fig. 1. Dynamics and shape distortion of reconstruction with Lorenz (x,z) variables for embedding dimensions 3 (a) and 4 (b). The values of dynamics distortion are sorted to increasing order (black line) and the values of shape distortion are plotted with regard to the sorting of the dynamics distortion measure (grey line).

5. Simulations and results

We compare by means of simulations the proposed embedding scheme to two other methods. These are the uniform univariate embedding (UU), i.e. delay embedding on a single time series disregarding the presence of other time series, and the extension of the uniform embedding to the multivariate case (UM), i.e. for p time series the dimension of the embedding vectors is $M = mp$. Uniform embedding suggests that a fixed delay τ is used, which is selected by the time-delayed mutual information criterion. Note that τ is separately computed for each time series and their average is used in UM. The embedding dimension M is selected by the FNN procedure with threshold 0.01. The FNN procedure is used in the same way as in the univariate case, but at each step the embedding vectors are augmented by the addition of the p components corresponding to the lagged multivariate time series. For the proposed non-uniform multivariate embedding (N-UM), the same threshold is used for the stopping criterion.

Noise-free bivariate time series as well as time series contaminated with 5% normal observational noise are generated from chaotic systems. The data are scaled to be in the same range [0,1] to compensate for different variable ranges. For each system we generate 100 time series of length $N = 8192$. For each realization we obtain the state space reconstruction with the three methods and compute Sd and Dd.

The bivariate time series are from the x and z variables of the Lorenz system

$$\dot{x} = 10(y - x), \quad \dot{y} = 28x - y - xz, \quad \dot{z} = -8/3z + xy,$$

the x and y variables of the Rössler system

$$\dot{x} = -y - z, \quad \dot{y} = x + 0.15y, \quad \dot{z} = 0.2 + xz - 10z,$$

and the y_1 and y_2 variables of the coupled Rössler–Lorenz system at two regimes: independent ($C = 0$) and strong coupling ($C = 4$)

$$\dot{x}_1 = 6(-y_1 - z_1), \quad \dot{y}_1 = 6(x_1 + 2y_1), \quad \dot{z}_1 = 6(0.2 + x_1 z_1 - 5.7z_1)$$
$$\dot{x}_2 = 10(y_2 - x_2), \quad \dot{y}_2 = 28x_2 - y_2 - x_2 z_2 + Cy_1^2, \quad \dot{z}_2 = -8/3z_2 + x_2 y_2.$$

For each bivariate time series, we consider the uniform and non-uniform multivariate embeddings, UM and N-UM, as well as the uniform univariate embedding from the first and second variable, UU1 and UU2, respectively.

In Table 1 we give the results for the shape distortion measure for each reconstruction scheme and each system, where we highlight the best reconstruction scheme for each system. We see that both multivariate embedding procedures preserve better the shape of the attractor in comparison to the univariate embeddings, as expected. Uniform and non-uniform multivariate embeddings give equivalent values for the shape distortion measure with uniform being slightly better for most of the systems.

In Table 2, the results on the dynamics distortion measure are presented in the same way. Again multivariate embedding procedures preserve better the dynamics of the original system, and uniform and non-uniform embeddings give similar results. For systems where the observed variables behave differently, i.e., Lorenz and uncoupled Rössler–Lorenz system, non-uniform embedding performs marginally better, while uniform embedding performs marginally better when the variables have similar behavior, i.e., Rössler x and y and strongly coupled Rössler–Lorenz.

Finally, in Table 3 we see the average embedding dimension obtained from the 100 simulations, i.e. the average number of components included in the embedding vector (a pair of values for the methods on the bivariate time series). The non-uniform multivariate embedding gives reconstruction vectors with smaller or equal dimension to these of the uniform multivariate embedding. Univariate embeddings mostly regard smaller dimensions of the embedding vectors because the full dynamics of the system may not be well represented in the observed variable. With reference to Table 2, N-UM has less dynamic distortion than UM whenever N-UM embedding dimension is less than that of UM. In general N-UM turns out to give reconstructions at state space of smaller dimensions without distorting the dynamics any more than UM does. This is particularly useful in applications where high-dimensionality may cause problems.

Table 1. Average of the shape distortion measure

Shape Distortion (Sd)	UU1	UU2	UM	N-UM
Lorenz (x, z)	0.647	0.357	0.805	**0.878**
Lorenz (x, z)+5% noise	0.154	0.102	**0.294**	0.278
Rössler (x, y)	0.797	0.813	**0.899**	0.869
Rössler (x, y)+ 5% noise	0.196	0.202	**0.382**	0.250
Coupled Rössler Lorenz $C = 0$ (y_1, y_2)	0.046	0.036	0.528	**0.586**
Coupled Rössler Lorenz $C = 0$ +5% noise	0.062	0.046	**0.474**	0.447
Coupled Rössler Lorenz $C = 4$	0.667	0.713	**0.854**	0.832
Coupled Rössler Lorenz $C = 4$ +5% noise	0.301	0.377	**0.470**	0.459

Table 2. Average of the dynamics distortion measure

Dynamics Distortion (Dd)	UU1	UU2	UM	N-UM
Lorenz (x, z)	0.036	1.144	0.031	**0.030**
Lorenz (x, z)+5% noise	0.131	1.146	0.106	**0.099**
Rössler (x, y)	0.036	0.025	**0.024**	**0.024**
Rössler (x, y)+ 5% noise	0.115	0.111	**0.106**	0.107
Coupled Rössler Lorenz $C = 0$ (y_1, y_2)	0.950	1.048	0.246	**0.213**
Coupled Rössler Lorenz $C = 0$ +5% noise	0.972	1.038	**0.270**	0.274
Coupled Rössler Lorenz $C = 4$	0.092	0.381	**0.044**	0.051
Coupled Rössler Lorenz $C = 4$ +5% noise	0.166	0.419	**0.116**	0.120

Table 3. Average Embedding dimension

Average embedding dimension	UU1	UU2	UM	N-UM
Lorenz (x, z)	3	4.68	(2 , 2)	(2.03 , 1.44)
Lorenz (x, z)+5% noise	4.35	7.09	(3 , 3)	(2.07 , 3.44)
Rössler (x, y)	3	3	(2 , 2)	(2.4 , 1.12)
Rössler (x, y)+ 5% noise	4	4	(2 , 2)	(2 , 2)
Coupled Rössler Lorenz $C = 0$ (y_1, y_2)	3	3.6	(3.12 , 3.12)	(3.2 , 2)
Coupled Rössler Lorenz $C = 0$ +5% noise	4	4.7	(3.14 , 3.14)	(3.9 , 2)
Coupled Rössler Lorenz $C = 4$	3	3	(2 , 2)	(2.9 , 1)
Coupled Rössler Lorenz $C = 4$ +5% noise	4	4.9	(2 , 2)	(3 , 1)

6. Conclusions

It is evident that multivariate embedding significantly improves attractor reconstruction when compared to univariate embedding. Both multivariate approaches, the standard uniform approach and the non-uniform approach we introduce here, attain equivalently good state space reconstruction as evaluated by the shape and dynamics distortion measures. As shown by the average embedding dimensions, the proposed method is more parsimonious and gives embedding at lower dimension. Due to data size limitation, the implementation of the proposed method requires simplifications that result in a sub-optimal solution to the embedding problem as set by the conditions for component-wise independence and best correlation to future states. Further investigation is needed towards optimal solution that would allow a more fair comparison to the standard uniform multivariate embedding, which matches directly the criteria for optimum embedding (minimum mutual information for the delay parameter and insignificant percentage of false nearest neighbors for the embedding dimension).

Acknowledgments

This paper is part of the 03ED748 research project, implemented within the framework of the "Reinforcement Programme of Human Research Manpower" (PENED) and co-financed at 90% by National and Community Funds (25% from the Greek Ministry of Development–General Secretariat of Research and Technology and 75% from E.U.–European Social Fund) and at 10% by Rikshospitalet, Norway.

References

1. F. Takens, *Lecture Notes in Mathematics* **898**, 365 (1981).
2. T. Sauer, J. A. Yorke and M. Casdagli, *Journal of Statistical Physics* **65**, 579 (1991).
3. H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, 1997).
4. J. Guckenheimer and G. Buzyna, *Physical Review Letters* **51**, 1438 (1983).
5. M. Paluš, I. Dvořak and I. David, *Physica A* **185**, 433 (1992).
6. R. Hegger and T. Schreiber, *Physics Letters A* **170**, 305 (1992).
7. D. Prichard and J. Theiler, *Physical Review Letters* **73**, 951 (1994).
8. H. D. I. Abarbanel, T. A. Carroll, L. M. Pecora, J. J. Sidorowich and L. S. Tsimring, *Physical Review E* **49**, 1840 (1994).
9. L. Cao, A. Mees and K. Judd, *Physica D* **121**, 75 (1998).
10. S. P. Garcia and J. S. Almeida, *Physical Review E* **72**, 027205 (2005).
11. L. M. Pecora, L. Moniz, J. Nichols and T. L. Carroll, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **17**, 013110 (2007).
12. A. M. Fraser, *Physica D* **34**, 391 (1989).
13. A. M. Fraser, *IEEE Transactions on Information Theory* **35**, 245 (1989).
14. M. B. Kennel, R. Brown and H. D. I. Abarbanel, *Physical Review A* **45**, 3403 (1992).
15. A. M. Fraser and H. L. Swinney, *Physical Review A* **33**, 1134 (1986).
16. K. Judd and A. Mees, *Physica D* **120**, 273 (1998).
17. A. Papana and D. Kugiumtzis, *Nonlinear Phenomena in Complex Systems* **11**, 225 (2008).
18. J. Rissanen, *Automatica* **14**, 465 (1978).

19. N. F. Rulkov, M. M. Sushchik, L. S. Tsimring and H. D. I. Abarbanel, *Physical Review E* **51**, 980 (1995).