



Analysis and prediction of aperiodic hydrodynamic oscillatory time series by feed-forward neural networks, fuzzy logic, and a local nonlinear predictor

Pier Luigi Gentili, Hiroshi Gotoda, Milos Dolnik, and Irving R. Epstein

Citation: *Chaos: An Interdisciplinary Journal of Nonlinear Science* **25**, 013104 (2015); doi: 10.1063/1.4905458

View online: <http://dx.doi.org/10.1063/1.4905458>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/chaos/25/1?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

[Intrinsic excitability state of local neuronal population modulates signal propagation in feed-forward neural networks](#)

Chaos **25**, 043108 (2015); 10.1063/1.4917014

[Input vector optimization of feed-forward neural networks for fitting ab initio potential-energy databases](#)

J. Chem. Phys. **132**, 204103 (2010); 10.1063/1.3431624

Retraction: "Theoretical investigation of the dissociation dynamics of vibrationally excited vinyl bromide on an ab initio potential-energy surface obtained using modified novelty sampling and feed-forward neural networks" [*J. Chem. Phys.* **124**, 054321 (2006)]

J. Chem. Phys. **125**, 079901 (2006); 10.1063/1.2264051

[Theoretical investigation of the dissociation dynamics of vibrationally excited vinyl bromide on an ab initio potential-energy surface obtained using modified novelty sampling and feed-forward neural networks](#)

J. Chem. Phys. **124**, 054321 (2006); 10.1063/1.2162170

[Isotropic vector hysteresis modeling with feed-forward neural networks](#)

J. Appl. Phys. **91**, 8322 (2002); 10.1063/1.1456400

An advertisement for the journal 'Computing in Science & Engineering'. It shows several copies of the journal standing upright, each displaying a different scientific visualization on its cover, such as a hurricane and a complex energy landscape. The journal title 'Computing in SCIENCE & ENGINEERING' is prominently displayed in the bottom right corner of the advertisement area.

Analysis and prediction of aperiodic hydrodynamic oscillatory time series by feed-forward neural networks, fuzzy logic, and a local nonlinear predictor

Pier Luigi Gentili,^{1,a)} Hiroshi Gotoda,² Milos Dolnik,³ and Irving R. Epstein³

¹Department of Chemistry, Biology and Biotechnology, University of Perugia, 06123 Perugia, Italy

²Department of Mechanical Engineering, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu-shi, Shiga 525-8577, Japan

³Department of Chemistry, Brandeis University, Waltham, Massachusetts 02454-9110, USA

(Received 16 October 2014; accepted 22 December 2014; published online 6 January 2015)

Forecasting of aperiodic time series is a compelling challenge for science. In this work, we analyze aperiodic spectrophotometric data, proportional to the concentrations of two forms of a thermoreversible photochromic spiro-oxazine, that are generated when a cuvette containing a solution of the spiro-oxazine undergoes photoreaction and convection due to localized ultraviolet illumination. We construct the phase space for the system using Takens' theorem and we calculate the Lyapunov exponents and the correlation dimensions to ascertain the chaotic character of the time series. Finally, we predict the time series using three distinct methods: a feed-forward neural network, fuzzy logic, and a local nonlinear predictor. We compare the performances of these three methods. © 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4905458>]

The prediction of chaotic time series generated by chemo-hydrodynamic systems is a compelling challenge for current chemical physics and related branches of nonlinear science. In this work, we use Takens' theorem to analyze the aperiodic time series produced by a hydrodynamic photochemical system that transforms constant ultraviolet (UV) irradiation into an oscillating spectrophotometric signal. We then compare three methods—feed forward neural networks (FFNN), fuzzy logic (FL), and a local nonlinear predictor (LNP)—in utilizing the initial portion of a chaotic time series to predict the remainder of the series.

I. INTRODUCTION

Prediction of aperiodic time series is an intriguing challenge for science. Aperiodic time series may appear in at least three different situations. First, they may be associated with multivariate nonlinear deterministic systems whose dynamics is extremely sensitive to initial conditions. Alternatively, they may arise from the dynamics of non-deterministic natural complex systems that consist of many agents. Third, aperiodic time series can also be due to fluctuations or to errors associated with taking measurements. Fluid flows involving convection, modeled by the Navier-Stokes equation, are important examples of multivariate nonlinear deterministic systems generating aperiodic time traces. The prices of products in a financial market give rise to aperiodic time series generated by multi-agent complex systems whose governing laws are not known. In fact, irregular market fluctuations depend on macroeconomic forces and on a huge number of deals due to agents that act independently

of each other. Aperiodic time series generated by nonlinear deterministic systems with low levels of instrumental noise are chaotic; they differ fundamentally from the stochastic time series of multi-agent complex systems and from the noisy time series generated by uncorrelated errors in the measurement procedure.

One way to examine any aperiodic time series is to build a phase space in which to view the dynamics. A phase space can be constructed by applying Takens' embedding theorem¹ for both deterministic and stochastic systems. A deterministic chaotic time series differs from a stochastic one in two essential respects. First, its trajectory in its phase space generates a strange attractor with a fractal dimension. Second, two trajectories with nearby initial conditions diverge at an exponential rate characterized by the largest Lyapunov exponent.²

The beginning of “modern” time series prediction might be set at 1927, when Yule invented the autoregressive technique in order to predict the annual number of sunspots.³ For the half-century following Yule, the reigning paradigm remained that of linear models driven by noise.⁴ Linear time series models have two particularly desirable features; they can be understood in great detail, and they are straightforward to implement. However, they may be entirely inappropriate for even moderately complicated systems. Two crucial developments in aperiodic time series prediction occurred around 1980 with the general availability of powerful computers: the first development was state-space reconstruction by time-delay embedding; the second was the emergence of the field of machine learning, typified by neural networks.

In this work, we analyze and predict aperiodic spectrophotometric time series (A_t) generated by a hydrodynamic “photochemical oscillator”⁵ consisting of a solution of a thermo-reversible photochromic species, 1,3-dihydro-1,3,3-trimethyl-8'-nitro-spiro[2H-indole-2,3'-[3H]naphth[2,1-b][1,4]oxazine] (spiro-oxazine for short, or SpO). Solutions of SpO in a 1 cm path length fluorimetric quartz cuvette are

^{a)}Author to whom correspondence should be addressed. Electronic mail: pierluigi.gentili@unipg.it. Tel.: (+39) 075-585-5573. Fax: (+39) 075-585-5598.

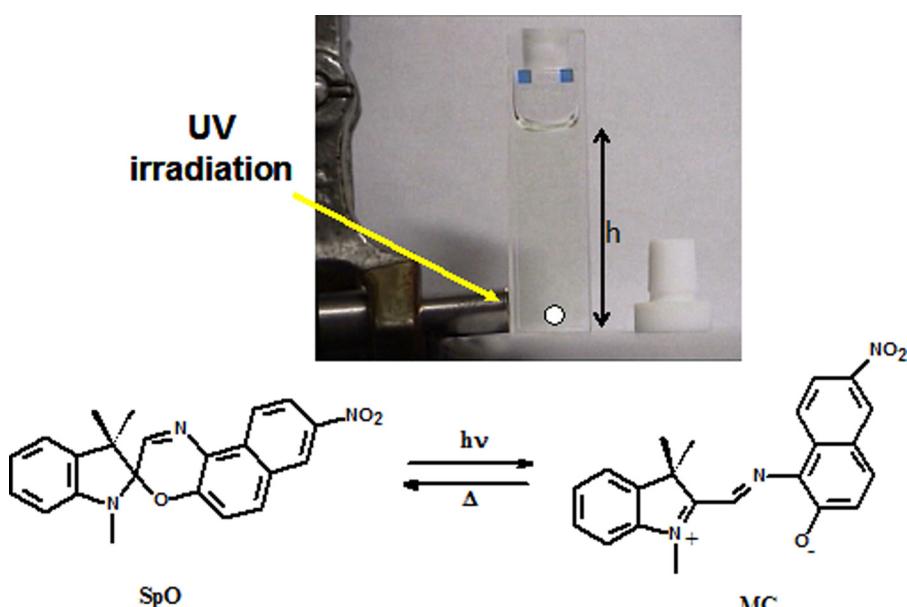


FIG. 1. Localized UV irradiation of a solution of a colorless spiro-oxazine (SpO) contained in an uncapped cuvette and producing a colored MC. The white spot at the front of the cuvette represents the beam of UV and visible light that probes how [MC] changes over time.

irradiated at the bottom with UV light (Figure 1). The light produces a colored photomerocyanine (MC) and at the same time heats the irradiated part of the solution. Experiments are performed in acetone and methanol as solvents, and in both capped and uncapped cuvettes. When the cuvette is uncapped, the partial evaporation of the solvent induces a cooling effect at the top (see Ref. 5 for more details on the extent of the cooling effect). The asymmetric irradiation of the solution and the vertical thermal gradient promote diffusive and convective motions. The chemistry of the photochromic species does not significantly affect the hydrodynamics, because it does not produce enough heat to cause substantial changes in the density of the solvent. On the other hand, the hydrodynamic flow causes more of the fresh solution containing colorless SpO to be irradiated with UV light. As a result of the hydrodynamic flow, the overall concentration of colored MC in the system increases. We record the time evolution of the absorbance in the UV and visible regions at the bottom of the cuvette (see white spot in Figure 1) with a spectrophotometer. The value of the absorbance in the visible is proportional to the concentration of MC, whereas that in the UV is mainly due to SpO, whose concentration is much higher than that of MC. We observe that such a system behaves like a “photochemical oscillator,” because it transforms constant UV irradiation into oscillating UV and visible transmitted intensities that are antiphase.⁵

A reaction-diffusion-convection model for this hydrodynamic photochemical oscillator has been developed in Ref. 5. In this paper, we present an analysis of the recorded spectrophotometric aperiodic oscillatory time series by building their phase space using Takens’ theorem, determining the Lyapunov exponents and the dimensions of their attractors. After demonstrating the chaotic character of the spectrophotometric time series, we show the results of their prediction obtained by three methods: (i) artificial FFNNs, (ii) FL, and (iii) a LNP. We compare their performances with the aim of finding the method that gives the best prediction. We also compare the computational effort for the three methods.

II. METHODS

A. Experimental methods

The details of our experimental methods can be found in our previous paper.⁵ In this work, we analyze four time series recorded under the four different conditions listed in Table I. All the time series were recorded with a cycle time of 3 s and for a time interval of at least 2550 s. The number, N , of data points for each time series is listed in Table I.

B. Computational methods

1. Phase space reconstruction

To analyze a time series containing N observations, $\mathbf{A}^* = \{A_i, i = 1, \dots, N\}$, we embed it in an m -dimensional phase space using the Takens time-delay embedding theorem.¹ The time series \mathbf{A}^* becomes a matrix whose elements are a time-lagged version of the original data

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} A_1 & A_{1+\tau} & \dots & A_{1+(m-1)\tau} \\ A_2 & A_{2+\tau} & \dots & A_{2+(m-1)\tau} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N-(m-1)\tau} & A_{N-(m-2)\tau} & \dots & A_N \end{pmatrix} \\ &= \begin{pmatrix} \bar{A}_1 \\ \bar{A}_2 \\ \vdots \\ \bar{A}_{N-(m-1)\tau} \end{pmatrix}. \end{aligned} \quad (1)$$

TABLE I. Experimental conditions for the four time series studied in this work.

| Time series | Solvent | Cuvette | N |
|-------------|----------|----------|------|
| (a) | Acetone | Uncapped | 956 |
| (b) | Acetone | Capped | 852 |
| (c) | Methanol | Uncapped | 1363 |
| (d) | Methanol | Capped | 1559 |

Each row vector of the matrix (1) is a single point in the phase space. The (integer) time delay τ is a dimensionless parameter defined as the actual time delay, τ_R , divided by the sampling interval, Δt . This parameter must be given before the minimum embedding dimension is determined. The number of observations, N , is much larger than both m and τ .

a. Time delay. The time delay τ must be chosen carefully. If τ is too small, then adjacent components A_i and $A_{i+\tau}$ will be too highly correlated for them to serve as independent coordinates. If τ is too large, then neighboring components are too uncorrelated for the purpose. To determine a suitable time delay, we use two methods: the autocorrelation and the mutual information methods, implemented in TISEAN code.⁶ The autocorrelation of \mathbf{A} is determined by using algorithm (2):

$$c(\tau) = \frac{1}{N} \sum_{i=1}^N A_i A_{i+\tau}. \quad (2)$$

The delay time is usually chosen² as the smallest τ value for which $c(\tau)/c(0) \leq e^{-1}$.

Mutual information⁷ is the other method suggested as an effective tool to determine a reasonable delay time τ . We need to compute

$$S(\tau) = - \sum_{ij} p_{ij}(\tau) \ln \frac{p_{ij}(\tau)}{p_i p_j}, \quad (3)$$

where for some partition on the real numbers, p_i is the probability of finding a time series value in the i th interval and $p_{ij}(\tau)$ is the joint probability that an observation falls in the i th interval and at the observation time τ later falls into the j th. A good candidate for a reasonable time delay is a marked minimum of $S(\tau)$. Unlike the autocorrelation, the mutual information method takes into account nonlinear correlations.

b. Embedding dimension. To determine the minimal embedding dimension m , we use two methods: the false nearest neighbor⁸ (implemented in TISEAN code) and Cao's method⁹ (implemented in MATLAB).

The method of the false nearest neighbor is based on the fact that choosing too low an embedding dimension results in points that are far apart in the original phase space being moved closer together in the reconstructed space. For each point \bar{A}_i in the m -dimensional phase space, the method looks for its nearest neighbor \bar{A}_j . It calculates the square of the Euclidean distance

$$R_{m,i}^2 = \sum_{k=0}^{m-1} (A_{i+k\tau} - A_{j+k\tau})^2. \quad (4)$$

In going from dimension m to dimension $m+1$ by time delay embedding, we add a $(m+1)$ th coordinate onto each of the vectors, \bar{A}_i and \bar{A}_j . After the addition of the new $(m+1)$ th coordinate, the distance between \bar{A}_i and \bar{A}_j is

$$R_{m+1,i}^2 = R_{m,i}^2 + (A_{i+m\tau} - A_{j+m\tau})^2. \quad (5)$$

We designate as a false neighbor any neighbor for which

$$\left(\frac{R_{m+1,i}^2 - R_{m,i}^2}{R_{m,i}^2} \right)^{1/2} = \frac{|A_{i+m\tau} - A_{j+m\tau}|}{R_{m,i}} > R_t, \quad (6)$$

where R_t is a given heuristic threshold. When the number of false neighbors is zero, m is the suitable embedding dimension. The problem is how to choose the threshold value R_t . To avoid this issue, Cao proposed another algorithm

$$E(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} \frac{R_{m+1,i}}{R_{m,i}}. \quad (7)$$

$E(m)$ is dependent only on the dimension m and the time delay τ . To investigate its variation from m to $m+1$, Cao defined $E1(m) = E(m+1)/E(m)$. $E1(m)$ stops changing when m is greater than a value m_0 , if the time series comes from an attractor. Then, $m_0 + 1$ is the best estimate of the minimum embedding dimension.

2. Discerning chaotic time series

An exponential divergence of initially nearby trajectories in phase space is an important indicator of the possible presence of deterministic chaos. Such divergence can be probed by determining the Lyapunov exponents. Moreover, dissipative systems that exhibit chaotic behavior often have an attractor in phase space that is strange. Strange attractors are typically characterized by a fractal dimensionality, D_c , which can be evaluated as the correlation dimension.

a. Lyapunov exponents. We calculate the maximal Lyapunov exponent by two methods based on Rosenstein's¹⁰ and Kantz's¹¹ algorithms.

Rosenstein's method locates the nearest neighbor of each point after reconstructing the phase space. The nearest neighbor, \bar{A}_j , is found by searching for the point that minimizes the initial distance to the particular reference point \bar{A}_i . This distance is expressed as

$$d_j(0) = \min_{\bar{A}_j} \|\bar{A}_i - \bar{A}_j\|, \quad (8)$$

where $\|\cdot\|$ denotes the Euclidean norm.

From the definition of the Lyapunov exponent, the j th pair of nearest neighbors diverge approximately at a rate given by the largest Lyapunov exponent after i discrete time steps (i.e., after a time $i\Delta t$, where Δt is the sampling period of the time series)

$$d_j(i) \approx d_j(0) e^{\lambda_1(i\Delta t)}. \quad (9)$$

In logarithmic form, Eq. (9) becomes

$$\ln d_j(i) \approx \ln d_j(0) + \lambda_1(i\Delta t). \quad (10)$$

The linear relation (10) can be determined for each point in the phase space. The largest Lyapunov exponent is calculated using a least-square fit to the “average” line defined by

$$y(i) = \frac{1}{\Delta t} \langle \ln d_j(i) \rangle, \quad (11)$$

where the average is carried out over all values of j .

Kantz's method computes

$$S(\Delta t) = \frac{1}{N} \sum_{j=1}^N \ln \left(\frac{1}{\|U_j\|} \sum_{i \in U_j} \text{dist}(T_j, T_i; \Delta t) \right), \quad (12)$$

where N is the length of the time series. The term $\text{dist}(T_j, T_i; \Delta t)$ is the distance between a reference trajectory T_j and a neighboring trajectory T_i after the relative time, Δt , determined as

$$\text{dist}(T_j, T_i; \Delta t) = \|\bar{A}_j - \bar{A}_i\|, \quad (13)$$

where \bar{A}_i is a neighbor of \bar{A}_j after Δt . U_j is the ε -neighborhood of \bar{A}_j , i.e., the set of all delay vectors of the series having a distance less than or equal to ε with respect to \bar{A}_j . $S(\Delta t)$ is calculated for different sizes of ε (ranging from a minimum given by data interval/1000 to a maximum given by data interval/100) and for different embedding dimensions. The maximal Lyapunov exponent is the slope of the curve $S(\Delta t)$.

b. Correlation dimension. The strangeness of an attractor can be measured by determining the correlation dimension D_c with the Grassberger-Procaccia algorithm,¹² which requires calculation of the correlation sum

$$C(m, \varepsilon) = \frac{1}{N'^2} \sum_{i \neq j}^{N'} \Theta(\varepsilon - \|\bar{A}_i - \bar{A}_j\|), \quad (14)$$

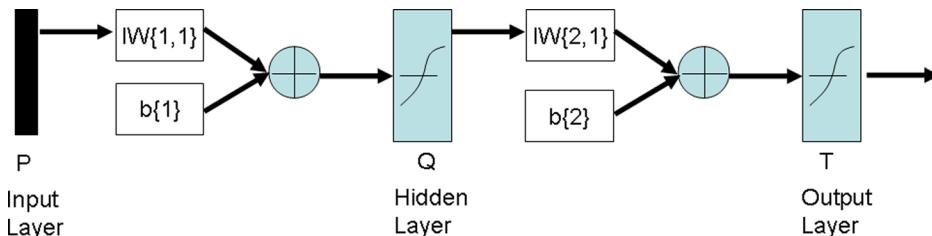
where ε is the size of the cell (whose minimum is given by data interval/1000 and whose maximum is the data interval), $N' = N - \tau(m-1)$, and Θ is the Heaviside step function, which is 1 when $\|\bar{A}_i - \bar{A}_j\| \leq \varepsilon$, whereas it is 0 when $\|\bar{A}_i - \bar{A}_j\| > \varepsilon$. Then, the correlation sum is exploited to measure the correlation dimension, assuming that

$$C(m, \varepsilon) \propto \varepsilon^{D_c}. \quad (15)$$

In practice, one checks for convergence of the estimated values of D_c as a function of the embedding dimension m .

c. Recurrence plots. The study of recurrences is used to understand the dynamics of nonlinear systems. One tool for such a purpose is the recurrence plot.¹³ A recurrence plot enables one to investigate the m -dimensional phase space trajectory through a two-dimensional representation of its recurrences. A recurrence plot is a two-dimensional square matrix with elements

$$R_{ij} = \Theta(\varepsilon - \|\bar{A}_i - \bar{A}_j\|), \quad (16)$$



where $\bar{A}_i \in R^m$ and $i, j = 1..N$. N is the number of considered states, \bar{A}_i ; ε is the threshold distance; $\|\cdot\|$ a norm and $\Theta(\cdot)$ the Heaviside function.

3. Predictive methods

There are many methods for predicting time series. In this work, we compare the performances of three methods: (i) a feed-forward back-propagation neural network; (ii) an adaptive neuro-fuzzy inference system (ANFIS); (iii) a local nonlinear predictor. All the kinetic data are divided into two parts: the first is used as a database (with N_{tr} points) for training the predictive method; a second part is used for the testing stage (with N_{ts} points).

a. Artificial neural networks. Artificial neural networks are algorithmization architectures simulating the behavior of networks of real nerve cells of the central nervous system. They are well suited to solving problems that are complex, ill-defined, and/or highly nonlinear, and to predicting chaotic and stochastic time traces. There are infinitely many ways to organize a neural network.^{14,15} In this work, we choose the most frequently used type of neural net, i.e., a feed-forward back-propagation network implemented through the Neural Network Toolbox of MATLAB. The feed-forward network we build consists of three layers: one input layer, one hidden layer, and one output layer. Figure 2 shows a schematic view of a feed-forward neural network with p neurons in the input layer and one bias, $b\{1\}$; one hidden layer with q neurons and one bias, $b\{2\}$, and finally one output layer with t neurons. A weight matrix ($IW\{1,1\}$) with dimension $[q \times p]$ is multiplied by a vector P , and the product is added to the bias $b\{1\}$ (which is a $[q \times 1]$ matrix); the result $IWP + b\{1\}$ forms the input to the hyperbolic tangent transfer function ($f = \text{tansig}$): $f(IWP + b\{1\}) = a\{1\}$. The output $a\{1\}$ of the hidden layer is weighted by the matrix $LW\{2,1\}$, which has dimension $[t \times q]$, and added to the bias $b\{2\}$ (a $[t \times 1]$ matrix). The sum, $LWa\{1\} + b\{2\}$, forms the input of a second hyperbolic tangent transfer function ($g = \text{tansig}$): $g(LWa\{1\} + b\{2\}) = a\{2\}$. The final output is $a\{2\}$.

The number of input and output neurons is set equal to the value of the embedding dimension, whereas the number of neurons in the hidden layer is fixed as $2m$. The matrix of data feeding the input layer is

$$\begin{pmatrix} A_1 & A_2 & \dots & A_{N_{tr}-(m-1)\tau} \\ A_{1+\tau} & A_{2+\tau} & \dots & A_{N_{tr}-(m-2)\tau} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1+(m-1)\tau} & A_{2+(m-1)\tau} & \dots & A_{N_{tr}} \end{pmatrix}. \quad (17a)$$

FIG. 2. Structure of the feed-forward neural network.

The matrix of data that constitutes the target of the training phase is

$$\begin{pmatrix} A_{1+T} & A_{2+T} & \dots & A_{N_{tr}-(m-1)\tau+T} \\ A_{1+\tau+T} & A_{2+\tau+T} & \dots & A_{N_{tr}-(m-2)\tau+T} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1+(m-1)\tau+T} & A_{2+(m-1)\tau+T} & \dots & A_{N_{tr}+T} \end{pmatrix}, \quad (17b)$$

where $T = 1$. The aim of the training phase is to adjust the weights of the connections so that the input pattern (17a) is mapped onto the corresponding output pattern (17b). In the testing phase, the matrix

$$\begin{pmatrix} A_{1s} & A_{2s} & \dots & A_{N_{ts}-(m-1)\tau} \\ A_{1s+\tau} & A_{2s+\tau} & \dots & A_{N_{ts}-(m-2)\tau} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1s+(m-1)\tau} & A_{2s+(m-1)\tau} & \dots & A_{N_{ts}} \end{pmatrix} \quad (17c)$$

is fed into the already trained network to predict the matrix (17d)

$$\begin{pmatrix} A_{1s+T} & A_{2s+T} & \dots & A_{N_{ts}-(m-1)\tau+T} \\ A_{1s+\tau+T} & A_{2s+\tau+T} & \dots & A_{N_{ts}-(m-2)\tau+T} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1s+(m-1)\tau+T} & A_{2s+(m-1)\tau+T} & \dots & A_{N_{ts}+T} \end{pmatrix}. \quad (17d)$$

The performance function for the feed-forward network is the mean square error (*mse*), that is, the average squared error between the network outputs a and the target outputs t :

$$mse = \frac{1}{N_{tr}} \sum_{i=1}^{N_{tr}} (t_i - a_i)^2. \quad (18)$$

The value of *mse* is back-propagated through the network, in the sense that the weights are modified according to their contribution to the error function. The final goal is to define the weights that minimize *mse*.

b. Fuzzy logic. Fuzzy logic models the human capability of computing with words: it is a rigorous logic of imprecise

and vague reasoning. It plays an important role in the development of artificial intelligence,^{16–19} and it is also processed by molecules.^{20–22} Fuzzy inference systems with adaptive capabilities are drawing increasing attention, since they can model the human ability to make decisions in complex situations and fine-tune the membership functions according to a desired input-output data set of any nonlinear function. In this work, we test the predictive power of ANFIS architectures using the Fuzzy Logic Toolbox of MATLAB. An example of the structure of an ANFIS is shown in Figure 3. It consists of five layers. The first contains the input variables; their number is equal to m , the value of the embedding dimension. The input variables feed the second layer, which consists of the input fuzzy sets. Each variable is partitioned into two fuzzy sets having generalized bell-shaped membership functions. A generalized bell-shaped membership function requires three parameters to be defined. The output of this layer is equal to the membership values of the input values to the fuzzy sets. The third layer contains the rules, which are of the Takagi-Sugeno type.²³ An example of such a rule is the following:

Rule k: If A_i is O_1 AND $A_{i+\tau}$ is B_2 AND $A_{i+2\tau}$ is C_1 AND $A_{i+3\tau}$ is D_1 AND $A_{i+(m-1)\tau}$ is E_1 , then $f_k = a_k A_i + b_k A_{i+\tau} + c_k A_{i+2\tau} + d_k A_{i+3\tau} + e_k A_{i+(m-1)\tau} + r_k$; where $O_1, O_2, B_1, B_2, C_1, C_2, D_1, D_2, E_1, E_2$ are the input fuzzy sets (Figure 3) and $a_k, b_k, c_k, d_k, e_k, r_k$ are the parameters of the k 'th rule.

For a set of input values, the absolute firing strength of each rule involving the AND operator is given by the product of the membership values of the input fuzzy sets involved in the rule

$$w_k = \mu_{O_1} \times \mu_{B_2} \times \mu_{C_1} \times \mu_{D_1} \times \mu_{E_1}. \quad (19)$$

The relative firing strength of the k 'th rule is the ratio of the k 'th rule's firing strength to the sum of all rules' firing strengths

$$\bar{w}_k = \frac{w_k}{\sum_j w_j}. \quad (20)$$

Layer 4 contains a number of nodes equal to the number of rules, and their outputs are given by

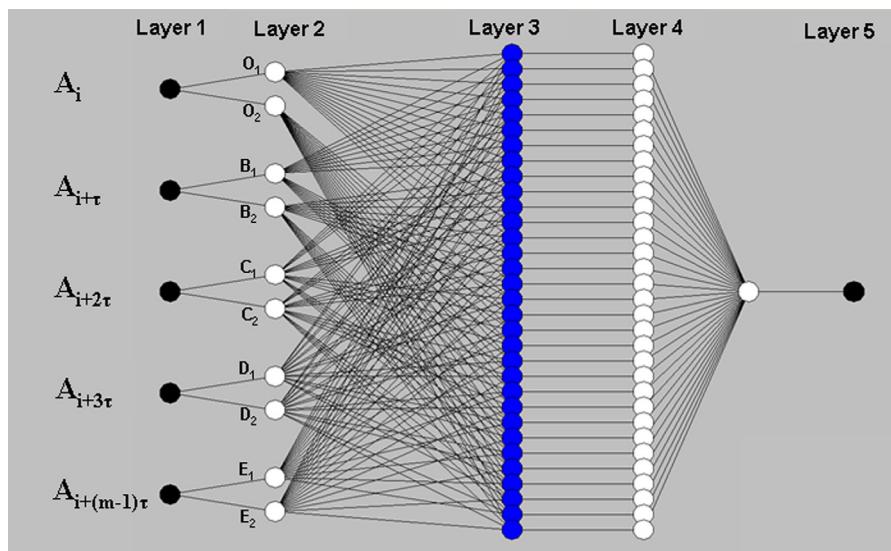


FIG. 3. Structure of an ANFIS with 5 input variables, each partitioned into two fuzzy sets, and involving 32 rules of the Takagi-Sugeno type.

$$\bar{w}_k f_k = \bar{w}_k [a_k A_i + b_k A_{i+\tau} + c_k A_{i+2\tau} + d_k A_{i+3\tau} + e_k A_{i+(m-1)\tau} + r_k]. \quad (21)$$

Finally, in layer 5 the defuzzification process occurs: the fuzzy output is transformed into a crisp output. In this work, we choose the weighted average (wtaver) method

$$\text{final output} = \sum_j \bar{w}_j f_j. \quad (22)$$

The matrix of data feeding the ANFIS is

$$\begin{pmatrix} A_1 & A_{1+\tau} & \dots & A_{1+(m-1)\tau} & A_{1+m\tau} \\ A_2 & A_{2+\tau} & \dots & A_{2+(m-1)\tau} & A_{2+m\tau} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots \\ A_{N_{tr}-m\tau} & A_{N_{tr}-(m-1)\tau} & \dots & A_{N_{tr}-\tau} & A_{N_{tr}} \end{pmatrix}. \quad (23)$$

The first m columns are the input vectors; the last column is the output vector. The training stage determines the best estimates of the $m \times 3 \times 2$ premise parameters (i.e., those regarding the input fuzzy sets) and the $n^{\circ}\text{rules} \times (m+1)$ consequent parameters. It is based on a hybrid method employing back propagation type gradient descent and least squares estimates.

c. Local nonlinear predictor. A proven nonlinear forecasting method has been proposed by Sugihara and May.²⁴ Its extended version, which includes the update of the library data, has also been extensively applied in a broad array of physical systems ranging from combustion and flame front instabilities^{25–27} to magneto-hydrodynamic instability.²⁸ In this work, we apply the original Sugihara-May method. As with the other predictive methods, the time series are divided into a training part, used as a source for generating library data, and a testing part. The nonlinear local predictor

requires that for each point \bar{A}_p of the trajectory in phase space, we search for its neighbors, that is, \bar{A}_k (with $k=1, 2, \dots, K$), where \bar{A}_p is the final point of trajectory in the phase space constructed from the data library. After T_s steps, \bar{A}_k is mapped to the T_s step-ahead prediction \bar{A}_{k+T_s} . The predicted \bar{A}_{p+T_s} is expressed as follows:

$$\bar{A}_{p+T_s} = \sum_{k=1}^K W(\bar{A}_k, \bar{A}_p) \cdot \bar{A}_{k+T_s}, \quad (24)$$

where W is a nonlinear weighting function depending only on the distances between \bar{A}_k and \bar{A}_p . We estimate the weighting function in a way similar to Refs. 24, 29, and 30.

$$\bar{A}_{p+T_s} = \frac{\sum_{k=1}^K \exp(-\|\bar{A}_p - \bar{A}_k\|) \cdot \bar{A}_{k+T_s}}{\sum_{k=1}^K \exp(-\|\bar{A}_p - \bar{A}_k\|)}. \quad (25)$$

Equation (25) is a local predictor in the sense that it refers to the K nearest library data. In this work, $K=10$.

4. Statistics for testing the quality of the predictions

To compare quantitatively the predictions achieved by the three methods, we use two statistics: the standard deviation σ_X and the correlation coefficient C_X . The standard deviation σ_X is defined as

$$\sigma_X = \left(\frac{1}{N_{ts}} \sum_{i=1}^{N_{ts}} (A_{i,\text{exp}} - A_{i,\text{pred}})^2 \right)^{1/2}, \quad (26)$$

where N_{ts} is the number of testing data, $A_{i,\text{exp}}$ is the i th experimental value, and $A_{i,\text{pred}}$ is the i th value predicted by the method. The correlation coefficient C_X is

$$C_X = \frac{N_{ts} \sum_i (A_{i,\text{exp}} A_{i,\text{pred}}) - \sum_i (A_{i,\text{exp}}) \sum_i (A_{i,\text{pred}})}{\left[\left(N_{ts} \sum_i (A_{i,\text{exp}}^2) - (\sum_i A_{i,\text{exp}})^2 \right) \left(N_{ts} \sum_i (A_{i,\text{pred}}^2) - (\sum_i A_{i,\text{pred}})^2 \right) \right]^{1/2}}. \quad (27)$$

The correlation coefficient always lies between $+1$ and -1 . When C_X is close to 1, there is a high degree of correlation. If C_X is negative, the data $A_{i,\text{exp}}$ and $A_{i,\text{pred}}$ are anti-correlated. When C_X is nearly zero, the predicted and the experimental data are independent, and the predictions are not reliable.

III. RESULTS AND DISCUSSION

A. Analysis of the time series

The spectrophotometric time trace recorded for the photochromic species dissolved in acetone and maintained in an uncapped cuvette is shown in Figure 4(a). Its Fourier

Transform (FT) spectrum is depicted in Figure 4(a'). The latter reveals that the time trace of Figure 4(a) is aperiodic. In fact, in the FT spectrum, there are two main families of frequencies (ν): one with peaks below 0.015 Hz and with a maximum at 0.0075 Hz (corresponding to a period of 133 s), and the other having peaks between 0.05 and 0.075 Hz with a maximum at 0.068 Hz (corresponding to a period of 15 s). When the same acetone solution is UV-irradiated in a capped cuvette, the thermal gradient achievable between the top and the bottom parts of the solution is lower, and the hydrodynamics is significantly different, as can be easily inferred by looking at the time trace of Figure 4(b). The amplitude of the oscillations is much smaller than before, and the behavior is

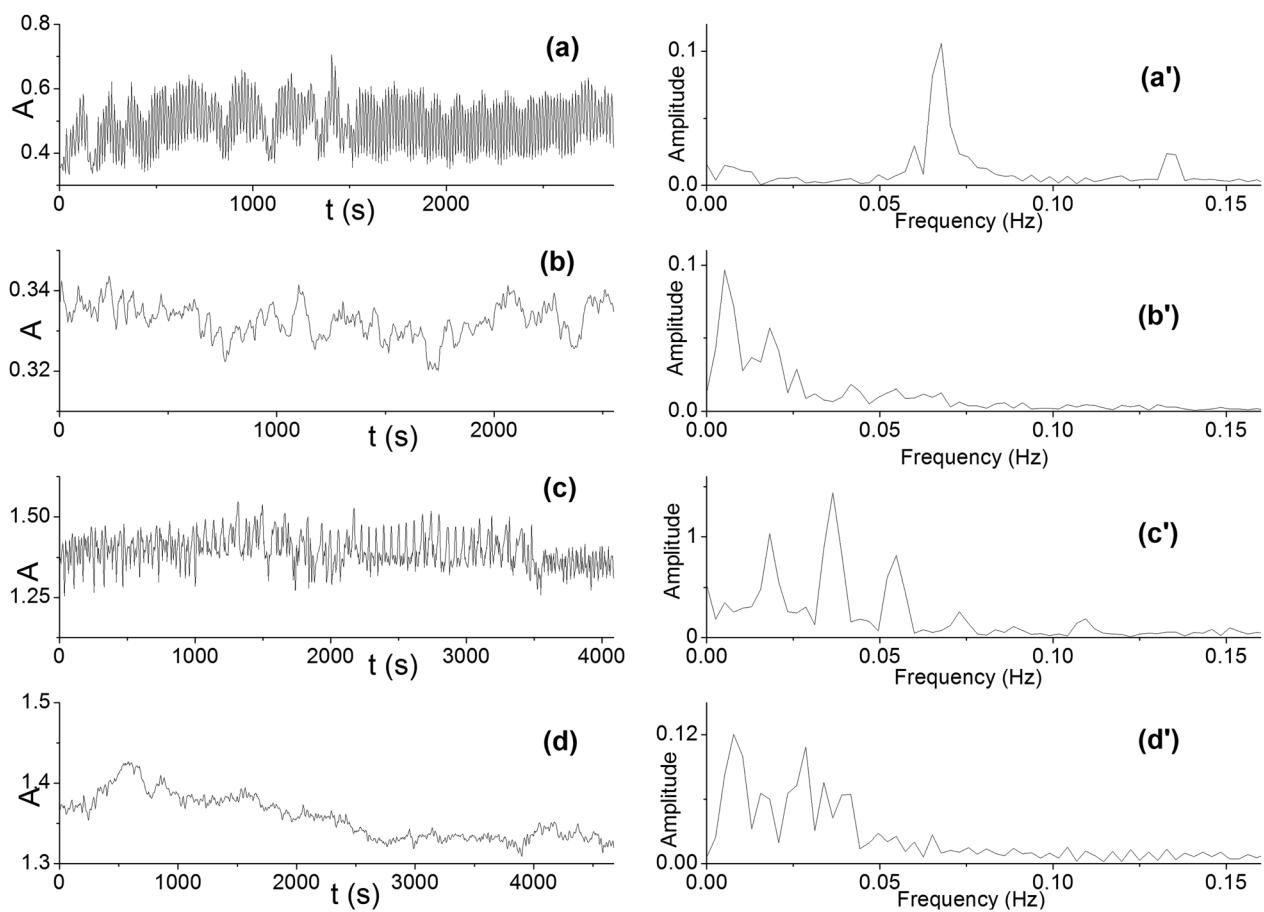


FIG. 4. Experimental time series of the absorbance (A) recorded at 612 nm in acetone and 607 nm in methanol and corresponding Fourier transforms (calculated in the time range 0–2553 s) for SpO dissolved in acetone in uncapped (a, a') and capped (b, b') cuvettes, and for SpO dissolved in methanol and maintained in uncapped (c, c') and capped (d, d') cuvettes.

much more irregular, as demonstrated by the FT spectrum in Figure 4(b'). The latter has appreciable components only at frequencies smaller than 0.024 Hz (corresponding to periods longer than 42 s).

The spectrophotometric time trace recorded for the photochromic species dissolved in methanol and maintained in an uncapped cuvette is plotted in Figure 4(c). The profile is aperiodic, as shown by the FT spectrum of Figure 4(c'). There is an array of components below 0.082 Hz, with three main peaks at 0.0176, 0.0296, and 0.0371 Hz (corresponding to periods of 57 s, 34 s, and 27 s, respectively). As in the case of acetone, with UV irradiation of the photochromic species dissolved in methanol in a capped cuvette, the time trace changes appreciably, exhibiting damped, more irregular oscillations (Figure 4(d)). The FT spectrum shows many

low-amplitude components with $\nu < 0.05$ Hz, losing the contribution of the larger ones.

Determination of the time delay through the autocorrelation and mutual information algorithms (see Sec. II B 1 a) reveals that τ is small for the (a) and (c) time series recorded with the cuvette uncapped, whereas it is large for the (b) and (d) time series recorded in a capped cuvette (Table II). In Figure 5(a), we plot the autocorrelation as a function of τ for the (a) time series. The embedding dimension determined by the false nearest neighbor and Cao's methods (see Sec. II B 1 b) is 5, except for the (d) time series, where $m = 4$. In Figure 5(b), we show the false nearest neighbors calculated for the (a) time series.

We calculate the largest Lyapunov exponent (λ) using the Rosenstein and Kantz algorithms (see Sec. II B 2 a) for the four time series. The Lyapunov exponent is positive in all four cases (Table II). We found a positive Lyapunov exponent also for time series simulated with our reaction-diffusion-convection model (see Ref. 5) of the photochemical oscillator in an uncapped cuvette with acetone as solvent. In Figure 5(c), we plot the results of Kantz's method applied to the (a) time series. Positive values of λ suggest that all the aperiodic time series are chaotic. This statement is corroborated by our determination of the correlation dimension D for the four time series by the Grassberger-Procaccia algorithm (see Sec. II B 2 b). The results are listed in Table II,

TABLE II. Best estimates of the time delay (τ) and its real value (τ_R), embedding dimension (m), Lyapunov exponent (λ), and correlation dimension (D_C) determined by the methods listed in Sec. II B.

| Time series | τ | τ_R (s) | m | λ (s^{-1}) | D_C |
|-------------|--------|--------------|-----|------------------------|-------|
| (a) | 2 | 6 | 5 | 0.064 | 3.37 |
| (b) | 17 | 51 | 5 | 0.00026 | 3.08 |
| (c) | 3 | 9 | 5 | 0.045 | 3.94 |
| (d) | 24 | 72 | 4 | 0.0074 | 3.18 |

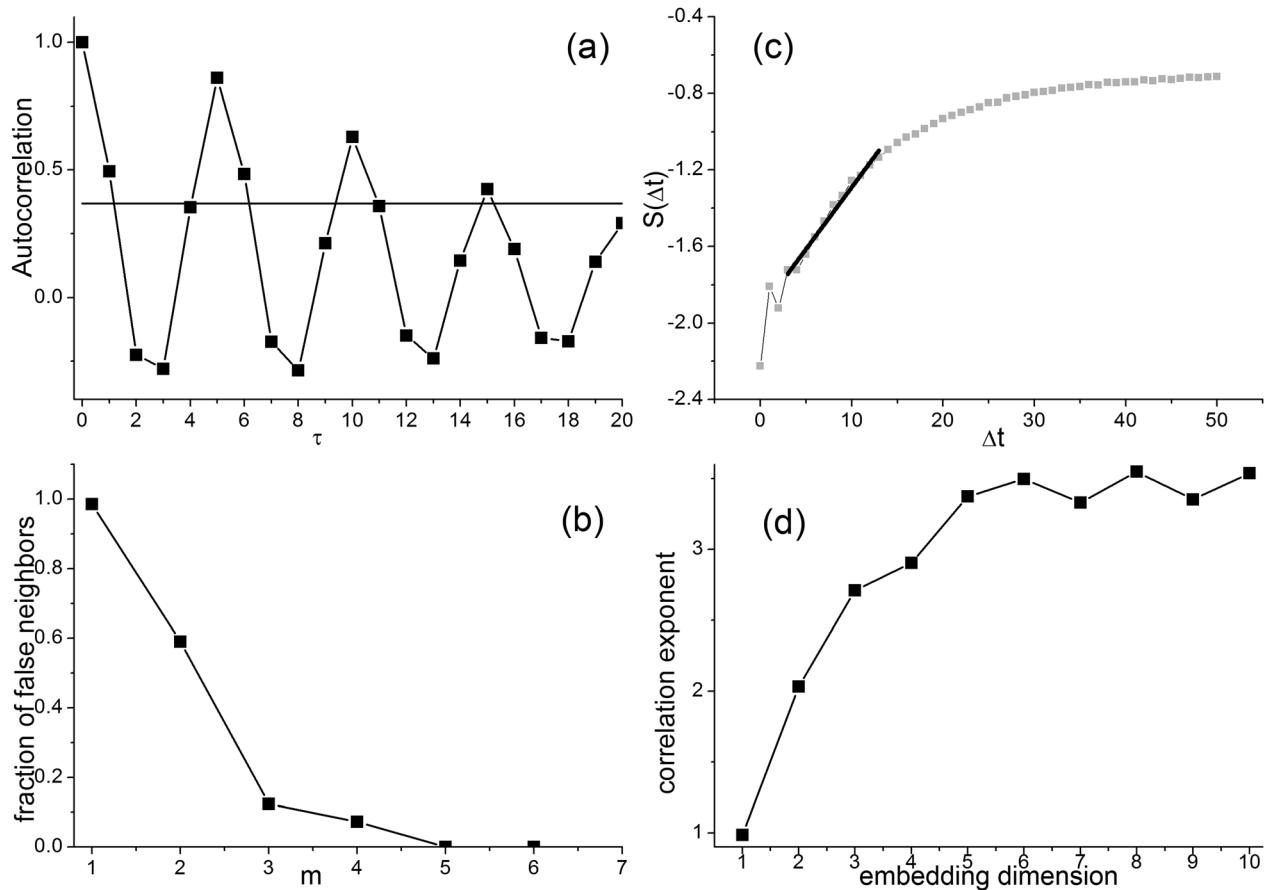


FIG. 5. Determination of the delay time by calculating the autocorrelation (a); determination of embedding dimension by calculating false nearest neighbors (b); determination of the Lyapunov exponent by Kantz's method (c), and determination of the correlation dimension (d). All plots utilize the (a) time series generated by SpO dissolved in acetone in an uncapped cuvette.

and a graphical example is shown in Figure 5(d) for the (a) time series.

The recurrence plots (see Sec. II B 2 c for details) of the four time series are shown in Figure 6. They allow a visual inspection of the high-dimensional phase space trajectories, because they trace with black dots the times at which a trajectory visits the same area in the phase space. Clearly, the four recurrence plots of Figure 6 have different typologies and textures.³¹ The plot derived from time series (a) shows a typology that is typical of a chaotic system, with diagonal, vertical, and horizontal lines. The recurrence plot of time series (c) is similar to (a), although it shows bright upper-left and lower-right corners, indicating nonstationarity but slowly varying parameters. A further reduction in the number of recurrent points is present in the plot of time series (b). Finally, the plot of time series (d) shows a disrupted typology with large white areas confirming conditions of nonstationarity.

The values of the two critical nonlinear invariants (λ and D) and the geometrical structure of the recurrence plots suggest that the aperiodic time series are chaotic. The chaotic character of our photochemical oscillator derives from a synergy between the thermo-reversible photochromic behavior of SpO and the convective dynamics of the solvent induced by the thermal gradient. Therefore, it is different from previous examples of deterministic chaos in chemistry like

bromate or chlorite oscillators. In fact, the earlier examples are pure chemical oscillators with no hydrodynamic convective motion.^{32,33}

The clear distinction between deterministic chaos and stochastic processes is a challenging topic. It is well-known that chaos has the important feature of short-term predictability and long-term unpredictability. Based on this property, one of us has recently proposed an extended version of the nonlinear local predictor as a useful distinction method between deterministic chaos and stochastic processes.²⁷ In this study, we focus on predicting the dynamics using three distinct methods. In future work, we plan to exploit the extended version of the local nonlinear predictor, including the permutation entropy in combination with a surrogate data method,³⁴ to obtain a deeper understanding of the nonlinear dynamics hidden in aperiodic spectrophotometric time series.

B. Prediction of the time series

After reconstructing the phase space, we tested the predictive power of our three methods: FFNN, FL, and LNP. For testing the quality of the predictions, we partitioned the four time traces into two data sets: the training and the testing data sets. The first of these serves as a database for training either the neural network or the fuzzy logic system, or

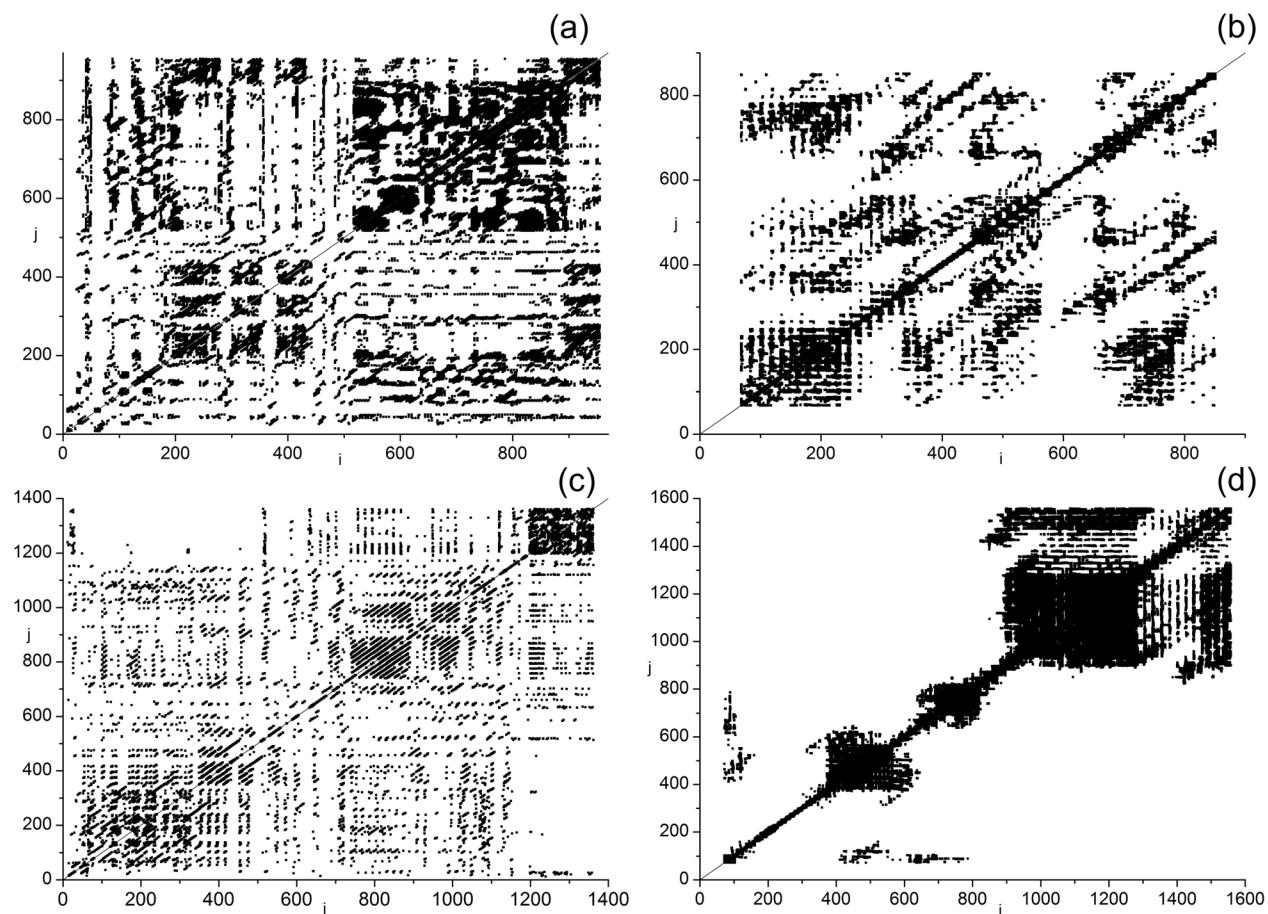


FIG. 6. Recurrence plots for SpO dissolved in acetone and maintained in an uncapped (a) and a capped cuvette (b), and for SpO dissolved in methanol and maintained in an uncapped (c) and a capped cuvette (d). The threshold distances ε are 0.05 in (a), 0.005 in (b), 0.03 in (c), and 0.01 in (d).

for searching the neighbors of the local predictor. The forecasting methods are supposed to find “rules” contained in the training set, and apply them to the points of the second part of the time trace. We tested the sensitivity of the predictions to the ratio (R) between the number of training data and the number of testing data by performing the calculations for R equal to 2 and 1.

In Figure 7, the testing data of time series (a) predicted by the FFNN, FL, and LNP methods for $R = 2$ are plotted as red traces in (1), (2), and (3), respectively. In the same graphs, the real testing data are shown as black traces. In (1'), (2'), and (3'), we report the prediction errors, calculated as $\Delta A = A_{\text{real}} - A_{\text{pred}}$. It is evident that the FFNN method with T (see Sec. II B 3 a) set to 1 is the most successful at predicting the chaotic trace. This is true also when $R = 1$. Both the fuzzy logic system based on five input variables and two fuzzy sets for each of them, and the local nonlinear predictor without updating give poorer results. In fact, the prediction error with FL and LNP is more than an order of magnitude larger than that of FFNN. This discrepancy is confirmed by the value of the standard deviations (σ_i , see Sec. II B 4 for its definition) and the correlation coefficients (C_i , Sec. II B 4) calculated considering only the testing data. The results are reported in Tables III and IV, respectively. The ratios of the standard deviations are $(\sigma_{\text{FL}}/\sigma_{\text{FFNN}}) \sim 2$ and $(\sigma_{\text{LNP}}/\sigma_{\text{FL}}) \sim 2$ at both $R = 2$ and $R = 1$. The FFNN and FL give positive and large correlation coefficient values, whereas LNP gives

slightly negative values. The FFNN gives the largest C_i , confirming the high correlation between its predicted and the real testing data set.

The results of the prediction of the testing data for time series (b) when $R = 2$ are depicted in Figure 8. Again, the FFNN method with T set to 1 is the most powerful predictor. In fact, C_{FFNN} is close to 1, whereas C_{LNP} is close to 0 and C_{FL} is negative. For $R = 1$, the standard deviations achieved with the three methods do not change; C_{FFNN} and C_{LNP} become slightly larger and C_{FL} remains negative.

For data set (c), the results when $R = 2$ are shown in Figure 9. Again, the FFNN method with $T = 1$ gives the best prediction, with $\sigma_{\text{FFNN}} = 0.023$ and $C_{\text{FFNN}} = 0.876$. For kinetics (c), the fuzzy logic system predicts the testing data set with positive C_{FL} , whereas the LNP predicts absorbance values that are anti-correlated with the real testing data (C_{LNP} is negative). For $R = 1$, the predictions made by the FFNN and LNP methods slightly worsen, whereas C_{FL} improves.

Finally, the results of the predictions for kinetics (d) when $R = 2$ are plotted in Figure 10. The FFNN method with $T = 1$ predicts the testing data quite well, more than three times better than a fuzzy logic system with eight bell-shaped fuzzy sets for four input variables ($\sigma_{\text{FL}}/\sigma_{\text{FFNN}} = 3.5$). The A values predicted by the LNP without updating (red trace in (3)) are always larger than the true A values (black trace in (3)) and σ_{LNP} is the largest among the three methods (Table

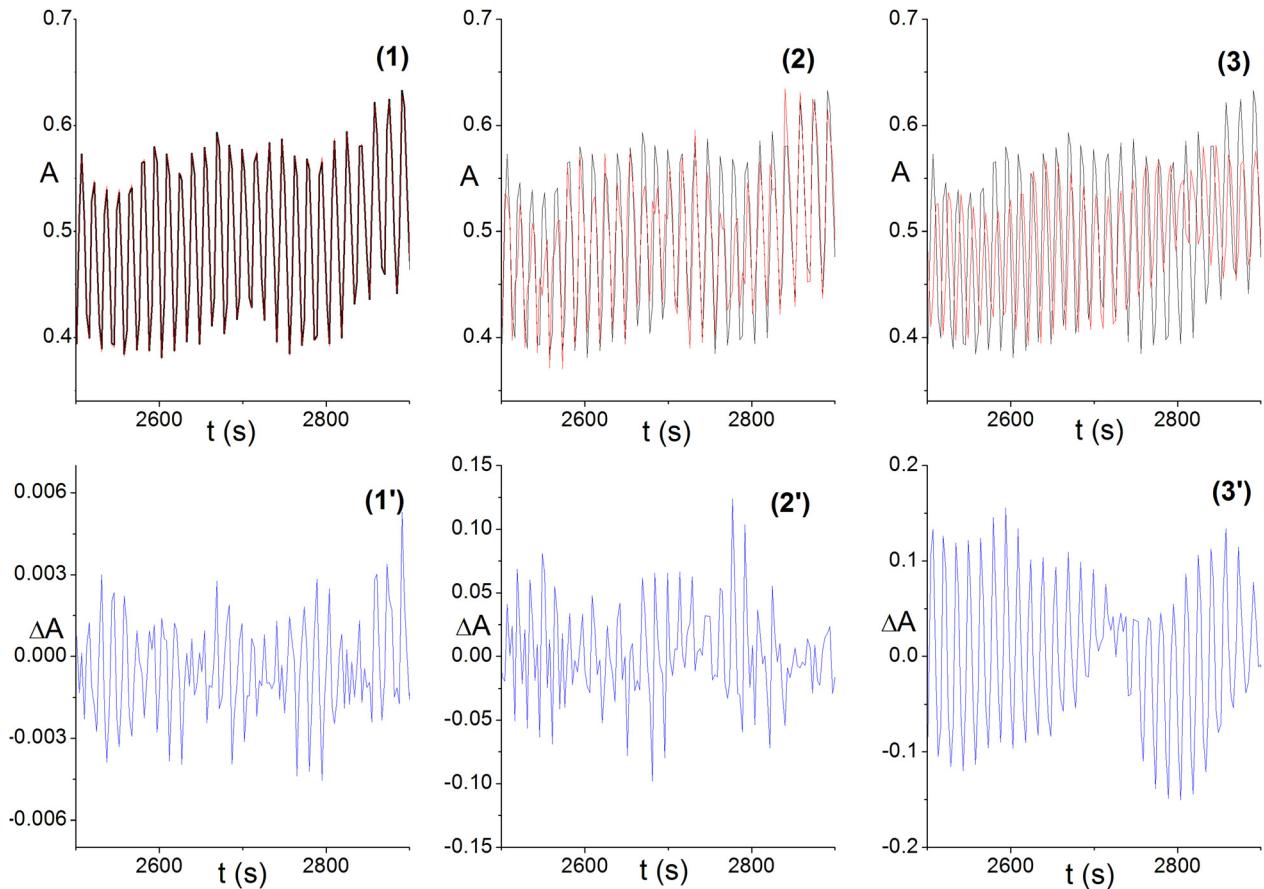


FIG. 7. Predictions (red traces) of the testing data set (black traces) for (a) kinetics when $R = 2$. The prediction errors are in blue. (1) and (1') refer to FFNN; (2) and (2') to FL; (3) and (3') to LNP.

IV). The performances of the FL and LNP methods become worse when $R = 1$, whereas that of FFNN does not change.

On the whole, the feed-forward neural network with $T = 1$ turned out to be the most effective method for predicting the aperiodic spectrophotometric traces. The predictions are generally more reliable when R is 2 than 1. The few exceptions are due to overfitting effects. In Table V, we report the time needed to perform the training stage for the three methods when $R = 2$. We clearly see that the LNP method was the most efficient in terms of computational effort. In fact, it required 10 parameters, which were determined in less than 0.016 s. On the other hand, the number of parameters required in the Feed-Forward Neural Network was 115 for kinetics (a), (b), (c), and 76 for (d); their determination occurred in tens of seconds. The Fuzzy Logic

Systems were the most expensive in terms of computational effort; they required the determination of 222 parameters for kinetics (a), (b), (c), 104 for (d) in times from tens to hundreds of seconds.

The performance of FFNN suffers drastically when T becomes large. For instance, the red trace in plot (1) of Figure 11 represents the prediction of the testing data of kinetics (a) when T is set to 10 and $R = 2$. The standard deviation is 0.037, more than two times larger than that obtained when $T = 1$ (Table III and Figure 7), and the correlation coefficient is 0.857, although the computational time did not change. This limitation of the FFNN has also been observed in the prediction of chaotic time series generated by nonlinear chemical reactions.³⁵ The predictive power of fuzzy logic may be improved by increasing the number of fuzzy sets. In

TABLE III. Standard deviations (σ) obtained for the testing data sets of four time traces with the three methods: FFNN, FL, and LNP with two data set partitions: $R = (N_{tr}/N_{ts}) = 2$ and 1.

| | (a) | (b) | (c) | (d) |
|-----------------|--|--|--|--|
| σ_{FFNN} | 0.016 ($R = 2$) 0.018 ($R = 1$) | 0.001 ($R = 2$) 0.001 ($R = 1$) | 0.023 ($R = 2$) 0.023 ($R = 1$) | 0.002 ($R = 2$) 0.002 ($R = 1$) |
| σ_{FL} | 0.030 ($R = 2$) 0.042 ($R = 1$) | 0.008 ($R = 2$) 0.008 ($R = 1$) | 0.061 ($R = 2$) 0.060 ($R = 1$) | 0.007 ($R = 2$) 0.072 ($R = 1$) |
| σ_{LNP} | 0.079 ($R = 2$) 0.085 ($R = 1$) | 0.005 ($R = 2$) 0.005 ($R = 1$) | 0.058 ($R = 2$) 0.059 ($R = 1$) | 0.058 ($R = 2$) 0.051 ($R = 1$) |

TABLE IV. Correlation coefficients (C) obtained for the testing data sets of four time traces with the three methods: FFNN, FL, and LNP with two data set partitions: $R = (N_{tr}/N_{ts}) = 2$ and 1.

| | (a) | (b) | (c) | (d) |
|------------|--|--|--|--|
| C_{FFNN} | 0.972 ($R = 2$) 0.967 ($R = 1$) | 0.913 ($R = 2$) 0.970 ($R = 1$) | 0.876 ($R = 2$) 0.872 ($R = 1$) | 0.972 ($R = 2$) 0.972 ($R = 1$) |
| C_{FL} | 0.897 ($R = 2$) 0.822 ($R = 1$) | -0.263 ($R = 2$) -0.360 ($R = 1$) | 0.230 ($R = 2$) 0.409 ($R = 1$) | 0.487 ($R = 2$) -0.117 ($R = 1$) |
| C_{LNP} | -0.073 ($R = 2$) -0.081 ($R = 1$) | 0.045 ($R = 2$) 0.164 ($R = 1$) | -0.017 ($R = 2$) -0.058 ($R = 1$) | -0.257 ($R = 2$) -0.358 ($R = 1$) |

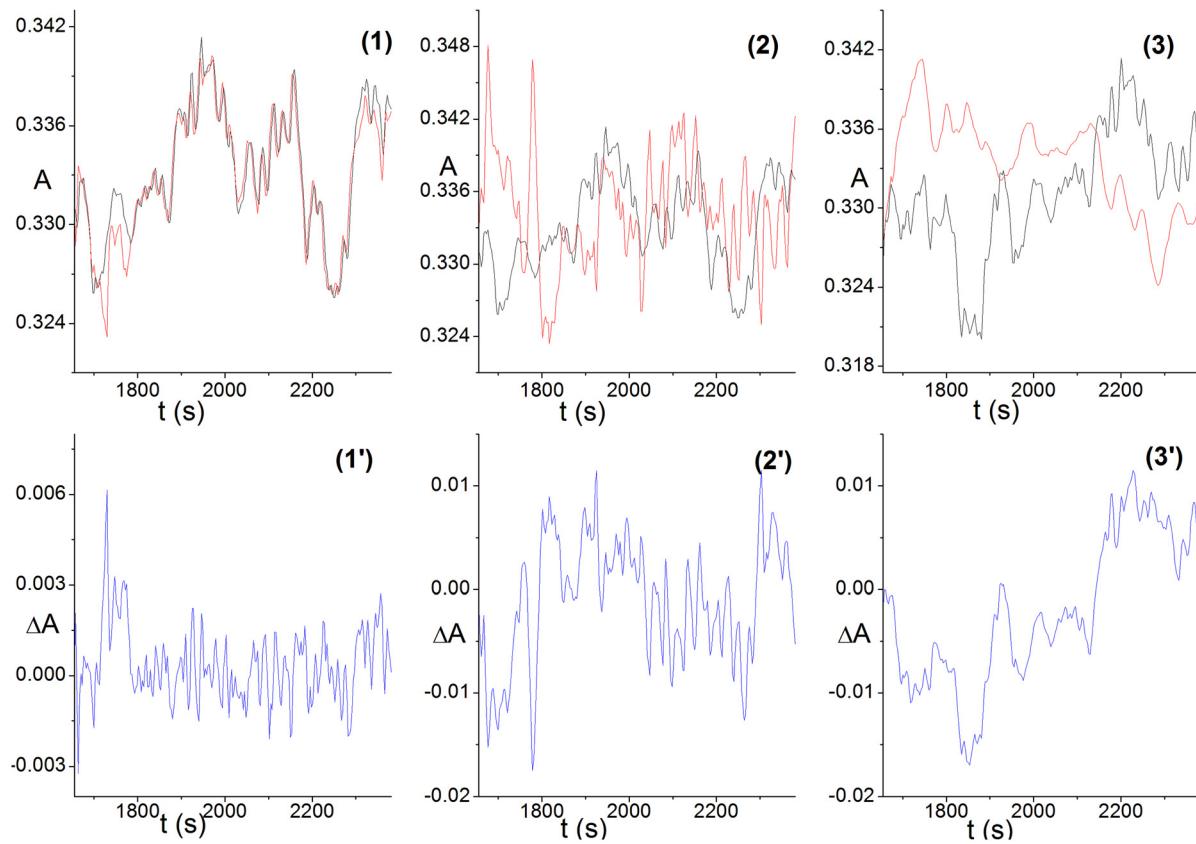


FIG. 8. Predictions (red traces) of the testing data set (black traces) for kinetics (b) when $R = 2$. The prediction errors are in blue. (1) and (1') refer to FFNN; (2) and (2') to FL; (3) and (3') to LNP.

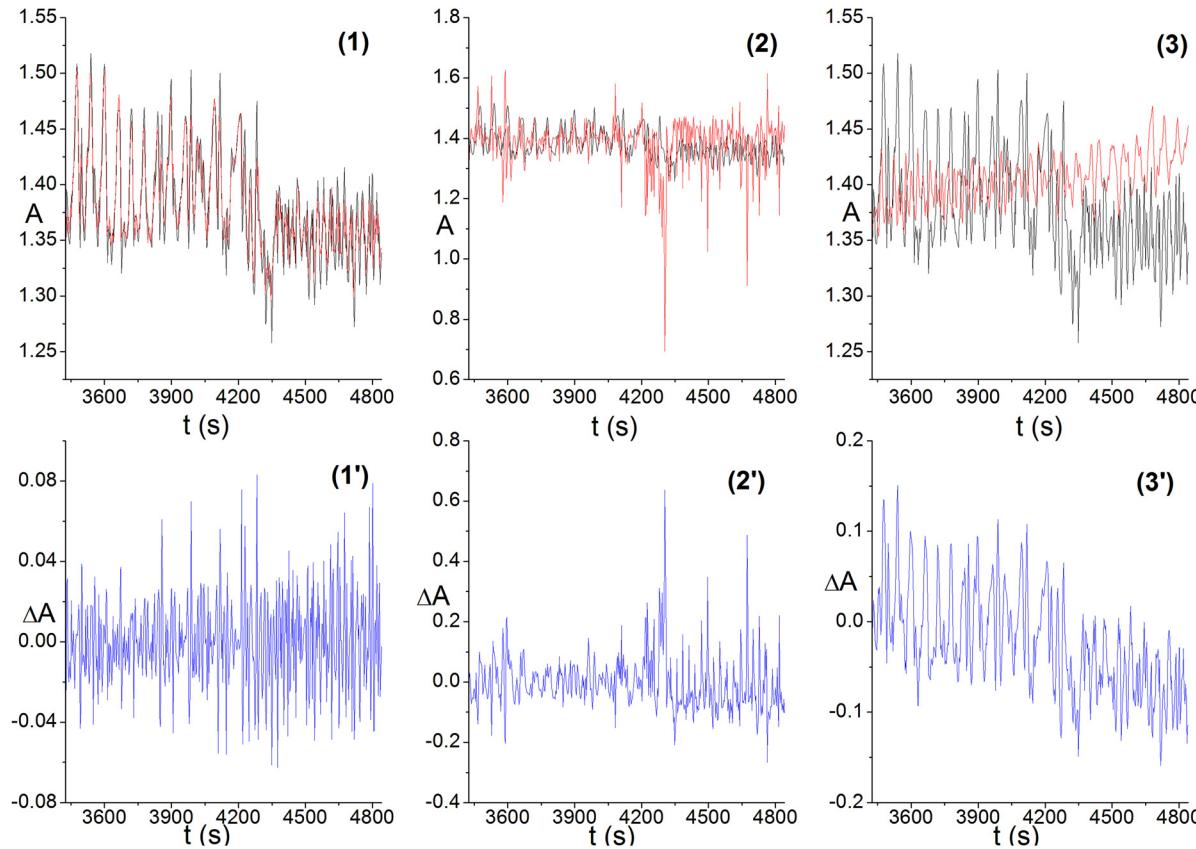


FIG. 9. Predictions (red traces) of the testing data set (black traces) for kinetics (c) when $R = 2$. The prediction errors are in blue. (1) and (1') refer to FFNN; (2) and (2') to FL; (3) and (3') to LNP.

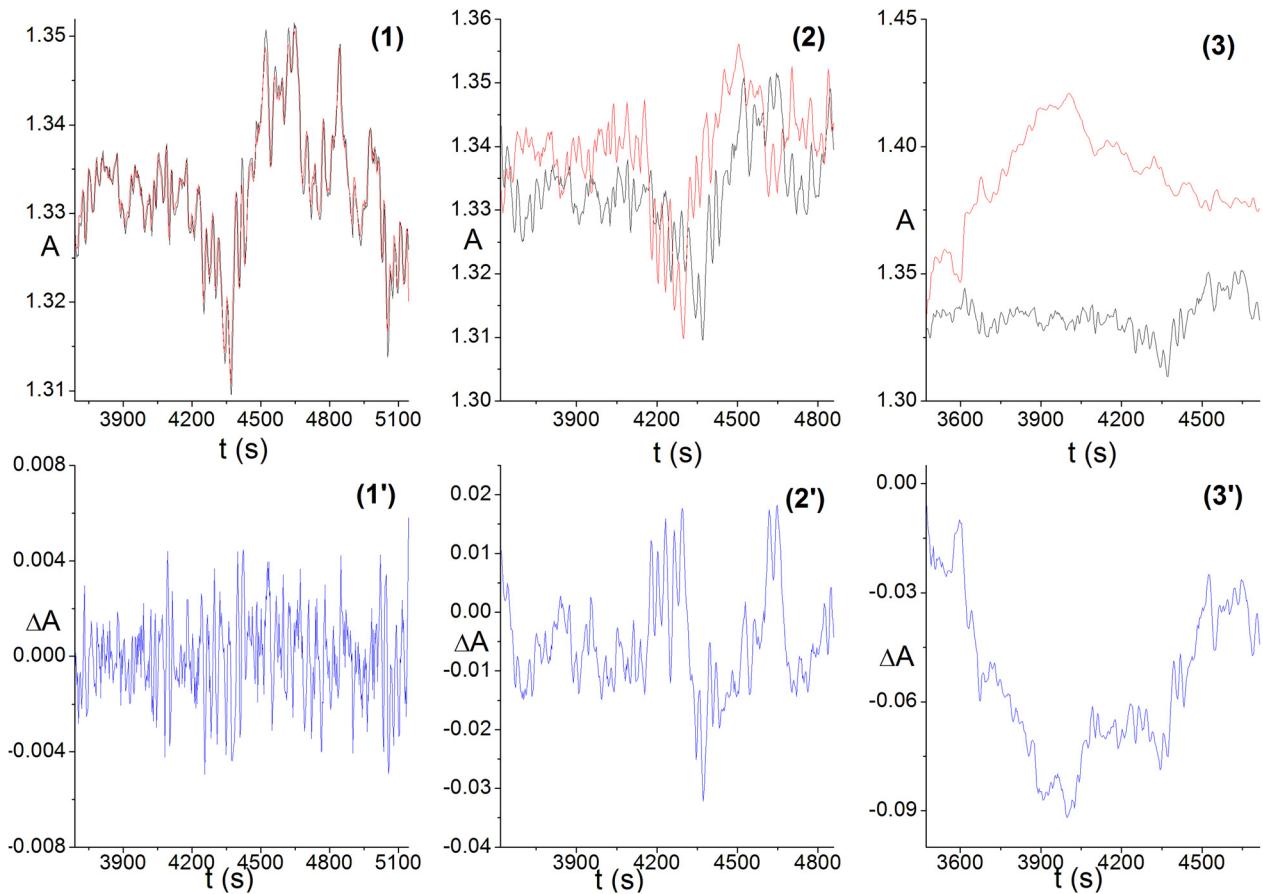


FIG. 10. Predictions (red traces) of the testing data set (black traces) for kinetics (d). The prediction errors are in blue. (1) and (1') refers to FFNN; (2) and (2') to FL; (3) and (3') to LNP.

plot (2) of Figure 11, the red trace represents the prediction achieved by a fuzzy logic system having four input variables and three bell-shaped fuzzy sets for each of them. The final standard deviation was equal to 0.020, almost two times smaller than when we used ten input fuzzy sets. The correlation coefficient also improved (to 0.958), at the expense of a greater computational effort (the time spent for 100 epochs was almost five times longer, i.e., 660 s). The forecasting performance of the local nonlinear predictor improves significantly when we update the library (see plots (3) and (3') in Figure 11). In fact, the standard deviation of the errors is almost four times smaller with updating the library than without (0.020 vs. 0.079), and the correlation coefficient value of 0.960 shows that the prediction becomes strongly correlated with the original data compared to the result of -0.073 obtained without updating. This significant improvement was achieved with a negligible increase in computational effort. In fact, the training stage required 0.14 s vs.

0.015 s spent without updating the library. When we update the LNP, we in effect transform the two stages, training and testing, of the predictive procedure into a single stage of constant training. On the other hand, our enhancements of the FFNN and FL preserve the distinction between the two stages. Through the training stage, we teach the neural net or the fuzzy logic system, and we define their parameters. Then, these parameters are not changed in the testing stage and are exploited to make predictions. The training stage thus requires a good “mapping” of the strange attractor. This is achieved with a small T in the case of the FFNN, and with a large number of fuzzy sets in the case of FL. The improvement of the LNP by updating adopts a different philosophy.

IV. CONCLUSIONS

Hydrodynamic photochemical oscillators consisting of solutions of SpO dissolved in organic solvents give rise to chaotic spectrophotometric traces, as demonstrated by their positive Lyapunov exponents and attractors with fractal dimensions. When we partition the matrix of data in two sets, that of training data and that of testing data, the best predictions of the chaotic profiles are achieved with a feed-forward neural network with T set to 1 and when the ratio between the number of training data and that of testing data is equal to two. The performance of the FFNN deteriorates when T is larger than 1. Fuzzy logic is a good predictive method only when we increase the number of input fuzzy

TABLE V. Time spent (in seconds) by our computing machine (Intel Core 2 Duo E6600) to complete the training process when $R = 2$. The number of steps or epochs to complete the training stage is 500 for FFNN, 100 for FLS.

| | (a) | (b) | (c) | (d) |
|------|-------|-------------------|-------|-------|
| FFNN | 36 | 24 | 44 | 23 |
| FLS | 146 | 96 | 191 | 42 |
| LNP | 0.015 | <10 ⁻⁶ | 0.016 | 0.015 |

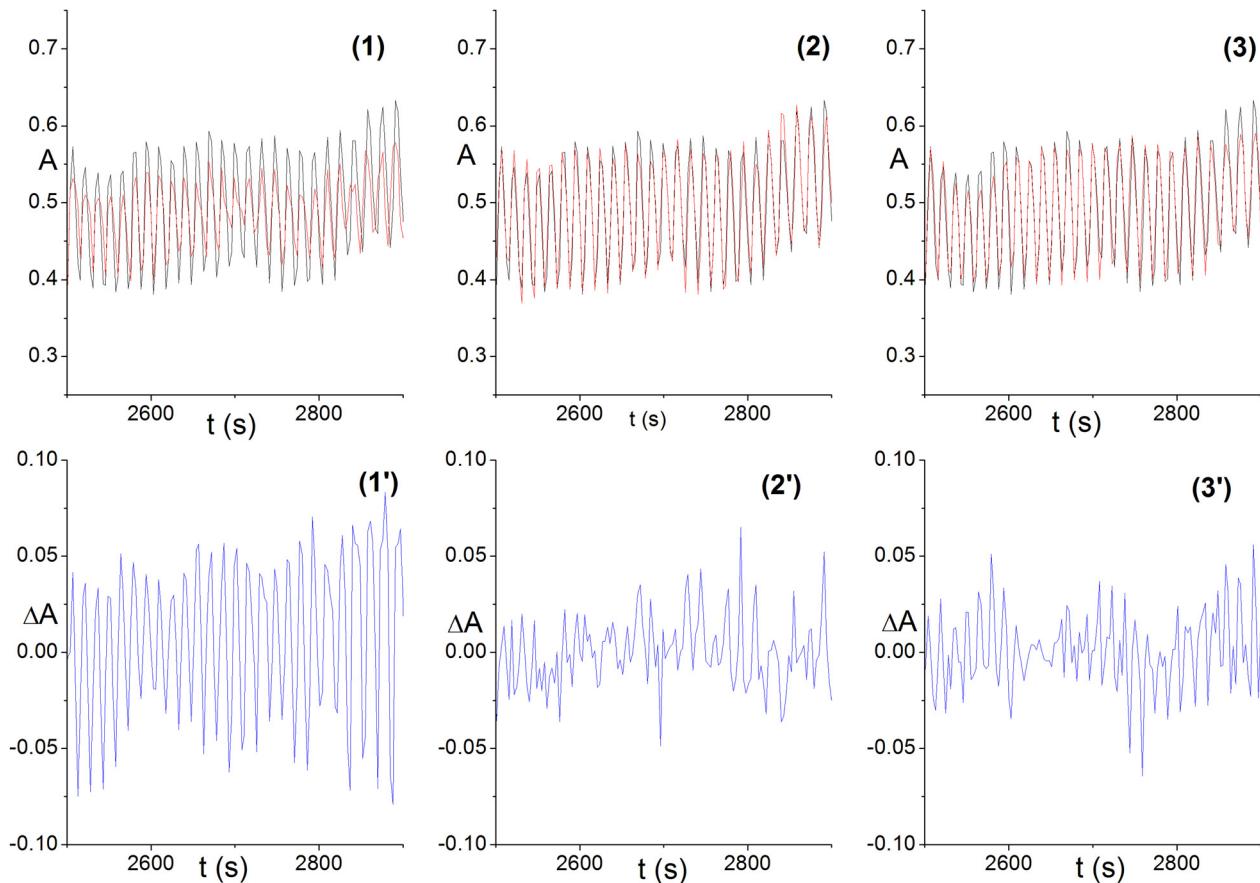


FIG. 11. Predictions (red traces) of the testing data set (black traces) for kinetics (a). The prediction errors are in blue. (1) and (1') refer to FFNN with $T = 10$; (2) and (2') to FL with three bell-shaped fuzzy sets for each variable; (3) and (3') to LNP with updating.

sets. However, this improvement requires a higher computational effort. Finally, the local nonlinear predictor is the cheapest method in terms of computational costs; however, it is only good at forecasting the chaotic kinetics when we update the library, thereby blurring the distinction between the training and testing data sets. In future work, we plan to seek a deeper understanding of the nonlinear dynamics hidden in our aperiodic time series and test the performance of other predictive methods such as other neural networks and genetic algorithms.

ACKNOWLEDGMENTS

P.L.G. acknowledges the financial support of the Ministero per l'Università e la Ricerca Scientifica e Tecnologica (Rome, Italy) and the University of Perugia [PRIN 2010-2011, 2010FM738P]. P.L.G. also acknowledges Dr. Angelo Facchini for providing preliminary instructions about the use of TISEAN code. M.D. and I.R.E. thank the U.S. National Science Foundation for support through grants CHE-1012428 and CHE-1362477.

¹F. Takens, “Detecting strange attractors in turbulence,” in *Dynamical Systems and Turbulence*, edited by D. A. Rand and L. S. Young (Springer, Berlin, 1981).

²Y.-C. Lai and N. Ye, *Int. J. Bifurc. Chaos* **13**, 1383 (2003).

³G. Yule, *Philos. Trans. R. Soc. London, Ser. A* **226**, 267 (1927).

⁴N. A. Gershenfeld and A. S. Weigend, “The future of time series: Learning and understanding,” in *Times Series Prediction: Forecasting the Future and Understanding the Past*, edited by A. S. Weigend and N. A. Gershenfeld (Addison-Wesley, Reading, MA, 1993).

⁵P. L. Gentili, M. Dolnik, and I. R. Epstein, *J. Phys. Chem. C* **118**, 598 (2014).

⁶R. Hegger, H. Kantz, and T. Schreiber, *Chaos* **9**, 413 (1999).

⁷A. M. Fraser and H. L. Swinney, *Phys. Rev. A* **33**(2), 1134 (1986).

⁸M. B. Kennel, R. Brown, and H. D. I. Abarbanel, *Phys. Rev. A* **45**, 3403 (1992).

⁹L. Cao, *Physica D* **110**, 43 (1997).

¹⁰M. T. Rosenstein, J. J. Collins, and C. J. De Luca, *Physica D* **65**, 117 (1993).

¹¹H. Kantz, *Phys. Lett. A* **185**, 77 (1994).

¹²P. Grassberger and I. Procaccia, *Phys. Rev. Lett.* **50**, 346 (1983).

¹³N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, *Phys. Rep.* **438**, 237 (2007).

¹⁴M. M. Nelson and W. T. Illingworth, *A Practical Guide to Neural Nets* (Addison-Wesley Publishing Company, Inc., USA, 1991).

¹⁵L. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications* (Prentice Hall, Upper Saddle River, NJ, 1994).

¹⁶L. A. Zadeh, *AI Mag.* **22**, 73 (2001).

¹⁷P. L. Gentili, *J. Intel. Fuzzy Syst.* **27**, 2137 (2014).

¹⁸P. L. Gentili, *RSC Adv.* **3**, 25523 (2013).

¹⁹D. Dubois and H. Prade, *Soft Comput.* **2**, 7 (1998).

²⁰P. L. Gentili, *Dyes Pigments* **110**, 235 (2014).

²¹P. L. Gentili, *ChemPhysChem* **12**, 739 (2011).

²²P. L. Gentili, *Phys. Chem. Chem. Phys.* **13**, 20335 (2011).

²³M. Sugeno and T. Yasukawa, *IEEE Trans. Fuzzy Syst.* **1**, 7 (1993).

²⁴G. Sugihara and R. M. May, *Nature* **344**, 734 (1990).

²⁵H. Gotoda, H. Nikimoto, T. Miyano, and S. Tachibana, *Chaos* **21**, 013124 (2011).

²⁶H. Gotoda, T. Ikawa, K. Maki, and T. Miyano, *Chaos* **22**, 033106 (2012).

²⁷H. Gotoda, Y. Shinoda, M. Kobayashi, Y. Okuno, and S. Tachibana, *Phys. Rev. E* **89**, 022910 (2014).

²⁸H. Gotoda, R. Takeuchi, Y. Okuno, and T. Miyano, *J. Appl. Phys.* **113**, 124902 (2013).

- ²⁹T. Miyano, H. Morita, A. Shintani, T. Kanda, and M. Hourai, *J. Appl. Phys.* **76**, 2681–2693 (1994).
- ³⁰H. Gotoda, T. Miyano, and I. G. Shepherd, *Phys. Rev. E* **81**, 026211 (2010).
- ³¹J.-P. Eckmann, S. Oliffson Kamphorst, and D. Ruelle, *Europhys. Lett.* **4**, 973 (1987).
- ³²K. B. Yatsimirskii and P. E. Strizhak, *Theor. Exp. Chem.* **28**, 293 (1992).

- ³³I. R. Epstein and J. A. Pojman, *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos* (Oxford University Press, 1998).
- ³⁴H. Gotoda, M. Amano, T. Miyano, T. Ikawa, K. Maki, and S. Tachibana, *Chaos* **22**, 043128 (2012).
- ³⁵D. Lebender and F. W. Schneider, *J. Phys. Chem.* **97**, 8764 (1993).