

## *Review Article*

# **Modeling Nonlinear Dynamics and Chaos: A Review**

**Luis A. Aguirre<sup>1</sup> and Christophe Letellier<sup>2</sup>**

<sup>1</sup> *Departamento de Engenharia Eletrônica da UFMG, Universidade Federal de Minas Gerais,  
Avenida Antônio Carlos 6627, 31270-901 Belo Horizonte, MG, Brazil*

<sup>2</sup> *CORIA UMR 6614, Université de Rouen, Avenue de l'Université, BP 12,  
76801 Saint-Etienne du Rouvray Cedex, France*

Correspondence should be addressed to Luis A. Aguirre, [aguirre@cpdee.ufmg.br](mailto:aguirre@cpdee.ufmg.br)

Received 28 January 2009; Accepted 24 February 2009

Recommended by Elbert E. Neher Macau

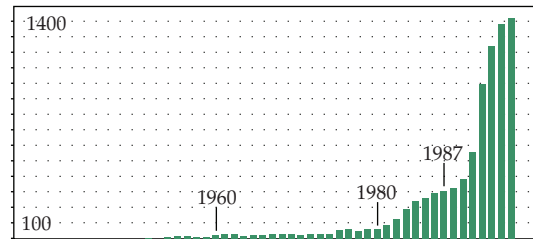
This paper reviews the major developments of modeling techniques applied to nonlinear dynamics and chaos. Model representations, parameter estimation techniques, data requirements, and model validation are some of the key topics that are covered in this paper, which surveys slightly over two decades since the pioneering papers on the subject appeared in the literature.

Copyright © 2009 L. A. Aguirre and C. Letellier. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## **1. Introduction**

The field of nonlinear dynamics experienced a very quick and intense development in the last thirty years or so. To determine the starting point of any subject is very difficult simply because we all stand on somebody else's shoulders in any "new" attempt we make. For the sake of presentation, and because we feel that this understanding is not absurd anyway, the origins of what we nowadays call the field of nonlinear dynamics can be traced back to the work of Henri Poincaré.

In his studies on nonlinear dynamical systems Poincaré figured out that, since no analytical solution to most of nonlinear systems can be obtained, the whole set of solutions can be investigated in the so-called phase space spanned by the set of variables required for a complete description of states of the system [1]. A few years later, while investigating the three-body problem, he observed that small perturbations can deeply affect the solution [2]. In order to investigate nonlinear dynamics, Poincaré introduced the concepts of phase portrait, Poincaré section, periodic orbit, return map, bifurcation, fixed point, and so on. Most of these concepts were used by Andronov's school in the 20s [3] but the first representation in the phase space of a now called "chaotic solution" was due to the solution Edward Norton Lorenz [4]. The so-called Lorenz attractor thus represents the first chaotic attractor ever drawn.



**Figure 1:** Number of papers published per year up to 1994. The search machine used was ISI Web of Science using Topic = (“nonlinear dynamics” OR chaos). Up to 1994 the total number of entries found was 8 400 (shown in this figure). Up to 2008 the total number of entries is 39 180 and only in 2008 over 2 700 papers were published within the topics nonlinear dynamics or chaos.

From the publication of Lorenz’s paper up to the mid of the 70 s not many papers were published. Among those who contributed to the emergence of “chaos theory,” we can quote Ruelle and Takens paper on turbulence [5], May about bifurcation [6], Rössler’s new chaotic attractors [7], Li and Yorke’s theorem for existence of chaos in map [8]. These contributions, published before 1980, popularized the word “chaos” as well as related techniques to investigate these new types of solutions.

An important turning point happened around 1980 (Figure 1). There can be little doubt that such a turning point was provoked, by the papers [9, 10]. The reason for this is quite simple. So far, in the investigation of a nonlinear system  $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ , with  $\mathbf{f} : \mathbb{R}^m \mapsto \mathbb{R}^m$ , typically it was assumed that the entire state vector  $\mathbf{x}$  was available. This constraint was clearly too restrictive in practical problems and consequently it was unclear what kind of role could the theory of nonlinear dynamics play in *real world problems*. From a suggestion by Ruelle, Packard and colleagues showed that it was possible to build a phase portrait for the original system using a scalar  $s(t) = h(\mathbf{x})$ , with  $h : \mathbb{R}^m \mapsto \mathbb{R}$  [9]. Furthermore, Takens and Mañé mathematically proved under which conditions the phase space reconstructed using  $s(t)$  was diffeomorphically equivalent to the one that would be obtained if the entire state vector  $\mathbf{x}$  was available [10, 11]. A decade later, other important works on the subject were published [12, 13]. When such works came to light, hopes of using the “new theory of nonlinear dynamics” in practical problems were kindled, because now all that was required, in principle, and assuming an ideal measurement function was the recording of a single variable.

When, in the early 1980, it became accepted that a scalar-reconstructed attractor could be equivalent to the original one, a great deal of work was devoted to developing tools to quantitatively characterize chaotic attractors. In the following decade, geometric measures such as dimensions [14, 15], Lyapunov exponents [16] and entropies [17] were adapted, applied, and intensively investigated in the context of nonlinear dynamics. By the end of the 1980, there was a large set of works on these subjects as can be found in [18, 19].

Another turning point which can be detected in the number of papers published in the field of nonlinear dynamics and chaos happened in the end of the decade 1980 (Figure 1). Unlike the turning point in the beginning of the decade, which quite clearly can be related to fundamental papers on embedding theory, it is quite hard to relate the turning point in the late 1980 to a specific topic within the field. Most likely, such a turning point was an indication that several important topics started to be investigated as well as many applications started in various fields. Among such topics we are to find the *modeling of nonlinear dynamics and chaos*. To survey the main developments of this topic is the aim of the present paper.

## 2. Nonlinear System Identification

Before actually addressing how the modeling of nonlinear dynamics and chaos developed within the field, it is important to say that a significant amount of work was developed independently, and sometimes previously, in the field of nonlinear system identification. The basic goals were very similar to those that later were pursued by the nonlinear dynamics community but there were some important differences. Nonlinear system identification, which has its origins in the field of engineering, is usually concerned with nonautonomous systems, discrete-time models, disturbance modeling (this is vital to correctly deal with noisy data), and hardly ever was concerned with chaos. On the other hand, modeling nonlinear dynamics and chaos, with its origins in physics and applied mathematics, usually concerned with autonomous systems, very often considers continuous-time models, does not typically model disturbances, and is strongly focused on chaotic systems.

Having established such main distinctions between the two fields, let us point out that this paper is a survey of modeling techniques applied to nonlinear dynamics and chaos. Many such techniques have indeed been developed in the field of nonlinear system identification and will be mentioned in this survey only to the extent in which they were applied to modeling nonlinear dynamics and chaos. In the remainder of this section, for the sake of completion, we present some issues in nonlinear system identification which will be relevant in the discussion of modeling techniques applied to nonlinear dynamic and chaos. In [20] the authors also point out some conceptual differences between what they call stochastic and deterministic modeling.

By the end of the 1970 linear system identification was well established. The typical problem was to build a model, usually a transfer function of the form

$$y(k) = \frac{B(q)}{A(q)}u(k), \quad (2.1)$$

from a set of possibly noisy data  $y(k)$ ,  $u(k)$ ,  $k = 1, 2, \dots, N$ . In (2.1)  $A(q)$  and  $B(q)$  are polynomials in the backward shift operator  $q^{-1}$ , that is,  $y(k-i) = q^{-i}y(k)$ . A commonly used representation of (2.1) is the ARX (autoregressive model with exogeneous inputs) model

$$y(k) = a_1y(k-1) + \dots + a_{n_y}y(k-n_y) + b_1u(k-1) + \dots + b_{n_u}u(k-n_u), \quad (2.2)$$

where  $n_y$  and  $n_u$  are the maximum lags used. Noise is usually assumed to appear in the output  $y(k)$ . The presence and the type of noise usually are such that the standard least squares estimator will produce wrong results. This calls for disturbance modeling and some type of generalized least squares algorithm, which is no longer a linear estimator. The simplest disturbance model which is the moving average (MA), which added to (2.2), yields the well-known ARMAX model. More general disturbance and system models have been proposed and studied in detail in standard texts such as [21, 22].

The approach to nonlinear system identification happened along different lines of action. One such line of action was to build so-called block-oriented models which consisted of a linear transfer function model (see (2.1)) either after or before a nonlinear *static* function  $g(\cdot)$ , such were the Hammerstein and Wiener models. Another line of action was the development of nonparametric models like the Volterra models [23]. The numerical algorithms involved in the development of such models were, in one way or another, related

to the correlation function. A glimpse of the state-of-the-art in the development of block-oriented and Volterra models at the end of 1970 can be found in [24].

A third, and last one to be mentioned here, line of action was to rewrite (2.2) as

$$y(k) = f(y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u)), \quad (2.3)$$

and to generalize for the case in which  $f(\cdot)$  was a nonlinear function. Because the regressor variables in (2.3) were of the ARX type, and because  $f(\cdot)$  was now allowed to be nonlinear, (2.3) was referred to as a NARX model. As before, in the case of noisy data, disturbance modeling is a must and, therefore, additional regressor terms of the MA (moving average) type have to be included in (2.3). This led to the now well-known acronym NARMAX. The building of a NARMAX model is typically composed of three steps: (1) the choice of the class of models to be used for approximating  $f(\cdot)$ . Some model classes include: polynomial models [25], rational models [26], neural models [27], radial basis function models [28], and wavelet networks [29]. (2) The following step is to determine which regressor variables to use in (2.3), this is equivalent, although it is more general, to the choice of the embedding space in modeling nonlinear dynamics and chaos. For recent discussions and a survey on such techniques we refer the reader to [30–35]. (3) In model building the final step is to estimate whatever parameters  $f(\cdot)$  may have. This step is far the easiest. Standard and robust techniques for accomplishing parameter are available [21, 22]. The important problem of model validation, which will be addressed later on, has more to do with model testing than model building.

### 3. Model Building for Nonlinear Dynamics and Chaos

After presenting a brief introduction (Section 1) and pointing out some standard problems in nonlinear system identification (Section 2), we are ready to start surveying the development of data-driven models for nonlinear dynamics and chaos. There are various possible ways of addressing this vast subject, and it is not clear which will turn out to be the most pedagogical one. We choose to start covering five of the earliest papers in the field and follow by discussing some of the mostly used model classes.

#### 3.1. A Few Pioneering Works

In what follows, we begin by mentioning explicitly five pioneering works, the first four of which were of great influence; [36–40]. The reason for including [40] in this group is its early date. In the remainder of the section we will survey some of the major developments in the field. Together, these five papers have been quoted 2719 times from their publication up to the end of 2008. (The search machine used was ISI Web of Science.)

##### 3.1.1. The Local Linear Predictor

The first papers on modeling nonlinear dynamics and chaos seem to have appeared in 1987. One of them is related to local linear modeling, the rest is generally concerned with some aspect of global modeling. Because of its historical importance, in this subsection, the main points concerning the local linear predictor [36, 41] will be reviewed.

Suppose that a measured time series  $y(1), y(2), \dots, y(N)$  lies on a  $D$ -dimensional attractor of an  $n$  th-order deterministic dynamical system. The starting point obtains an embedding from the recorded data. A convenient, though not unique, representation is achieved by using *delay coordinates*, for which a delay vector has the following form:

$$\mathbf{y}(k) = [y(k) \ y(k - \tau) \ \dots \ y(k - (d_e - 1)\tau)]^T, \quad (3.1)$$

where  $d_e$  is the *embedding dimension* and  $\tau$  is the *delay time*. Takens has shown that embeddings with  $d_e > 2n$  will be faithful generically so that there is a smooth map  $f : \mathbb{R}^{d_e} \mapsto \mathbb{R}$  such that

$$\mathbf{y}(k + 1) = f(\mathbf{y}(k)) \quad (3.2)$$

for all integers  $k$ , and where the forecasting time  $T$  and  $\tau$  are also assumed to be integers. Takens' theorem gives no indication as to how find. One of the first attempts to build the map  $f$  from the data was the local linear predictor method, which started by partitioning the embedding space into neighbourhoods  $\{\mathcal{U}_i\}_{i=1}^{N_n}$  within which the dynamics can be appropriately described by a linear map  $g : \mathbb{R}^{d_e} \mapsto \mathbb{R}$  such that

$$\mathbf{y}(k + 1) \approx g_i(\mathbf{y}(k)) \quad \text{for } \mathbf{y}(k) \in \mathcal{U}_i, \ i = 1, \dots, N_n. \quad (3.3)$$

Several choices for  $g$  have been suggested in the literature such as linear polynomials [36, 42] which can be interpolated to obtain an approximation of the map  $f$  [43]. Simpler choices include *zeroth-order approximations*, also known as *local constant predictors* [36, 44], and a *weighted predictor* [45]. Linsay proposed to weigh the predictions according to information obtained from the local data [45]. In [46] the authors put forward the concept of *pseudofalse neighbors* which are left out from the construction of a local linear predictor to improve performance, and in [47] the authors consider the multivariate case.

A common difficulty of such approaches is that the data have to be separated into neighbourhoods. Thus given a point in the embedded space the closest neighbours to such a point must be found. It is well-known that for many methods most of the CPU time is spent in searching for close neighbours in the embedding space within the data [48] and that the effort required to accomplish this grows exponentially with the embedding dimension.

Another drawback has to do with model representation. The local linear predictor does not have a closed form, in other words, it is not a global model. This prevents using the obtained model in any kind of analytical investigation.

The local linear predictor has been applied to several real-time series including Wolf's sunspot numbers [20, 49], daily dollar exchange rates [50], R-R heart intervals [49], to mention a few.

### 3.1.2. Equations of Motion

In 1987 James Crutchfield and Bruce McNamara published a paper that became a reference in the field [37]. The title of this section is a reference to that paper rather to a particular class of models. It seems safe to say that this was the first journal paper in the field devoted to global modeling, as opposed to the "patchwork-models" produced by the local linear approach, and which the authors call the "atlas equation of motion procedure."

In [37] the authors consider linear-in-the-parameter models of the type  $\dot{\vec{x}} = \vec{F}(\vec{x})$ , with  $\vec{F}(\vec{x}) = A^T \vec{\phi}(\vec{x})$ , where  $\vec{\phi}$  is a vector of function basis and  $A$  is a vector of parameters to be estimated from data. Discrete-time models are also considered in the paper. The authors described their method without any particular class of function basis in mind. The parameter vector was estimated using a singular value decomposition implementation of least squares. Their method is general and quite basic: the embedding dimension and the number of basis functions used are varied over a set of alternatives. The chosen model is the one with smallest entropy  $I(M)$ . In their examples, that include the Hénon map, Duffing, and van der Pol oscillators, the function basis used was B-splines and discrete monomials.

In closing, it should be said that more than a method, Crutchfield and McNamara put forward a general philosophy for modeling chaotic data. With the exception of some issues that are specific to such data, their procedure does not differ substantially from works in the field of nonlinear system identification using linear-in-the-parameter models widely available at that time (see [34] for a survey).

### 3.1.3. Multivariable Functional Interpolation

The paper by David Broomhead and David Lowe is among the four pioneering works that we have selected to start surveying the main developments in modeling of nonlinear dynamics and chaos [38]. Such a paper is often quoted as being responsible for the proposition of using Radial Basis Function (RBF) models to describe nonlinear systems. Although in their paper the authors quote a review work by M. J. D. Powell entitled “Radial basis functions for multivariable interpolation: a review,” it seems that the work by Powell did not consider dynamical systems but only static problems.

A typical model, as proposed by Broomhead and Lowe, can be written as

$$y(k) = b_0 + \sum_{i=1}^p \omega_i \phi(\|\mathbf{x}(k-1) - \mathbf{c}\|), \quad (3.4)$$

where  $p$  is the number of radial basis functions,  $\phi$ ,  $b_0$ , and  $\omega_i$  are the unknown parameters to be estimated,  $\mathbf{c}$  are the centers, and  $\|\cdot\|$  is usually the Euclidean norm. As it will be seen in the next section, Martin Casdagli also worked with RBF models, as pointed out by Broomhead and Lowe: “We note that in this application our approach is very close to that of Casdagli who has applied radial basis functions to the construction of nonlinear maps from time series data. (Here the authors quote Casdagli’s paper as “submitted to *Physica D*.” That paper, that will be mentioned in the next section, was eventually published in 1989.) Unlike Casdagli who used strict interpolation, we will employ the least squares generalization.” By “strict interpolation” the authors mean that the number of unknowns (parameters) and constraints (built from data) is the same (more on this in Section 3.1.4).

This paper is curious in some aspects. First of all, it is clearly devoted to machine learning for purposes of classification. Several pages are devoted to the problem of approximating the exclusive-OR function. The last examples in the paper seem to be an afterthought and consist of the modeling of the doubling map  $x_{n+1} = 2x_n$  (modulo 1) and of the quadratic map  $x_{n+1} = 4x_n(1 - x_n)$ . Also, only three works are listed in the references which are related to modeling nonlinear dynamics and chaos: [36, 39] and a preprint by Lapedes and Farber [51]. (Such a manuscript was at the time submitted to Proceedings of the IEEE. It seems that this paper was never published as such, but has circulated as a technical

report. This report is cited by a number of early papers on modeling nonlinear dynamics and chaos, proving to have been quite influential.) It is also very curious that no paper or result on embedding theory was used by Broomhead and Lowe. Because both their examples use 1D maps, there was no need for any serious embedding. In summary, the contribution of [38] seems to be the suggestion of using RBF models for nonlinear dynamics, but for the general view of the embedding-modeling problem we should look into [37] as the pioneering work. The fusion of such papers, that is, the use of RBF models in an embedding environment to model and predict nonlinear dynamics, happened in the next paper to be considered.

### 3.1.4. Prediction of Chaotic Time Series

From the abstract of the paper we clearly read the objective of the paper: “numerical techniques are presented for constructing nonlinear predictive models directly from time series data” [39]. The basic motivation for such was to “convincingly distinguish low-dimensional chaos from randomness.” In fact, Casdagli argues in favor of his model-based approach when comparing it with the use of dimension calculations and Lyapunov exponents to distinguish noise from chaos.

With respect to the model class, this paper is remarkable because it compared polynomial, rational, and RBF models, plus local linear predictions. Very few papers present results for more than one model class, a more recent exception to this rule is [52].

As for the use of “strict interpolation” (Broomhead and Lowe) or simply “interpolants” (Casdagli) or “approximants” (Casdagli), the situation is not completely clear. The paper mentions both. When describing general techniques (polynomial and rational models), Casdagli mentions the least squares solution. However, when presenting the RBF model the only case considered is that of choosing the model structure so as to have an inverse problem with unique solution. At the end of that section, the least squares solution is mentioned again, but as “ongoing research.” One of the critical problems of the “strict interpolation” approach is the total lack of robustness to noise. In fact, Casdagli does not consider any noise in the “global modeling” examples, but only in the “distinguishing noise from chaos” examples. In two occasions interesting remarks are made. In discussing modeling difficulties Casdagli says: “this does not appear to be due to numerical errors in the least-squares fitting and matrix inversion algorithms” [39, page 343]. This gives the impression that he was actually implementing both: the least-squares algorithm (to fit approximants) and the matrix inversion algorithm (to find the solution to the strict interpolation problem). The first was probably related to polynomial and rational models and the latter to RBF models. The second case, when discussing the use of RBF models for predicting invariant measures, the number of data points is quoted as  $N = 25$  and the data were noise-free. This clearly suggests “strict interpolation.”

The RBF models considered by Casdagli were of the form (compare with (3.4))

$$y(k) = \sum_{i=1}^{N-1} \omega_i \phi(\|\mathbf{x}(k-1) - \mathbf{c}\|) + \sum_{n=1}^{\hat{d}} \mu_n p_n(\mathbf{x}), \quad (3.5)$$

where  $N$  is the number of data iterates,  $p_n$  is a basis of the space of polynomials of degree at most  $d$  from  $\mathbb{R}^n$  to  $\mathbb{R}$ , with  $d$  known, and  $\mu_n$  are parameters. The author points out that “frequently the polynomial term is not included,” which is interesting in the light of other results, to be mentioned in Section 3.2.3.

Casdagli investigated the accuracy of short-term predictions based on model iteration and the *direct* method, which amounts to fitting a predictor with a single prediction horizon equivalent to the various one-step-ahead predictions of the *iterative* method. In the various cases investigated, the iterative method proved superior. A similar aim was pursued in the context of neural networks in [53], in the context of Volterra expansions in [54] and using a kernel-based approach in [43]. Finally, the author showed that the fitted RBF models were capable of reproducing some invariants, such as Poincaré sections and bifurcation diagrams. Although he only considered the noise-free case, a very significant step towards serious model validation was given. Perhaps the greatest lack in this relevant paper is that there is no mention on structure selection. Model instability under iteration is prematurely attributed to the model class and not to an unsuitable model structure.

By model class we mean the *type* of models used, as for instance polynomials, RBFs or neural networks. By model structure we mean the model *topology*, of a particular class. For instance, in the case of RBF models, the model structure is determined by the type and number of basis functions used, the lags used to compose the input space, and so on.

### 3.1.5. Construction of Differential Equations

The early paper by Cremers and Hübler had the clear objective to “obtain a concise description of an observed chaotic time sequence” [40]. From the beginning, the authors were clearly concerned with chaotic data. The model class they considered was a basis of Legendre polynomials, and in the introduction of the paper it was stated that the true objective was to estimate the parameters of a differential equation. This more restricted aim was confirmed later in [55].

The authors discuss a few aspects of embedding and the effect of dynamical noise. In their numerical examples no noise is considered. They mention successful estimation of the parameters for two systems: Lorenz and an autonomous version of the van der Pol oscillator, for which the following 4th-order approximation was used:

$$\dot{x}_j = \sum_{l,m=0}^{l+m \leq 4} c_{j,l,m} P^l(x_1) P^m(x_2), \quad j = 1, 2, \quad (3.6)$$

where  $c_{j,l,m}$  are the parameters to be estimated. The Legendre polynomials  $P^l(x_1)$  and  $P^m(x_2)$  must be evaluated at  $x_1$  and  $x_2$ , in addition to the time derivatives of  $x_1$  and  $x_2$  at each point in state space considered. The derivatives were estimated numerically, but there is no clear indication as to how this was accomplished nor what would be the effect on such estimates of measurement noise (which they considered not). The authors justify the use of a polynomial expansion by stating “if nothing is known about the properties of the flow vector field, a fit by a polynomial (sic) series is frequently favourable” [40, page 800]. This last remark results from a general theorem by Weierstrass stating that any analytical function can be approximated by an infinite polynomial expansion.

Breeden and Hübler later remarked that the 1987 paper also considered the discrete-time case, which is definitely not obvious. Then, so far the field of nonlinear dynamics is concerned, this paper seems to mark the beginning of building differential polynomial equations from data. Although no mention on important practical issues is made in this paper, such as structure selection, derivative and parameter estimation, it has the merit of having opened the way for a prosperous modeling class of techniques, to be surveyed in Section 3.2.1.



The citation relationships among these five pioneering papers are illustrated in Figure 2. In the center of the diagram the embedding papers [9, 10] appear as a common feature, with the exception of the paper [38] (see discussion in Section 3.1.3). This quotation diagram suggests that Cremers and Hübler were not aware of the other papers on model building. (Their paper was submitted in September 1986.) It is interesting to notice (not shown in the diagram) that Broomhead & Lowe and Casdagli mention the work of Lapedes and Farber (see [51]).

### 3.2. Model Classes

In the description to follow, the order of the model classes mentioned is somewhat arbitrary. We follow a rough chronological order of the first papers in each class.

#### 3.2.1. Continuous-Time Polynomials

A number of papers appeared in the early nineties which had as a common goal fitting ordinary differential equations (ODEs) to observed data. As pointed out in Section 3.1.5, the paper [40] was pioneering in this field, although it did not deal with any practical issue. Unlike Cremers and Hübler, most papers that followed, instead of using a basis of Legendre polynomials, used "Taylor-series expansions." As an example, such an expansion for a 2D system is [56]

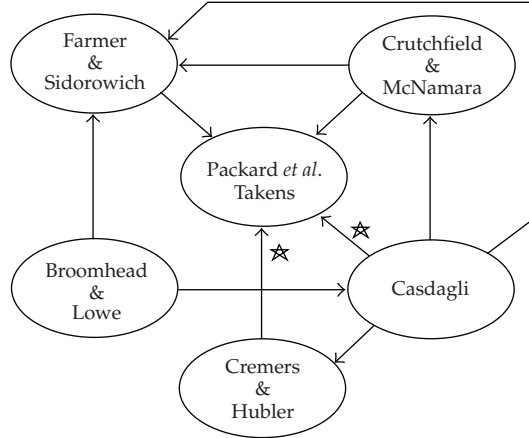
$$\begin{aligned}\dot{x} &= f(x, y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy + \cdots + a_{ij}x^i y^j + \cdots, \\ \dot{y} &= g(x, y) = b_{00} + b_{10}y + b_{01}x + b_{11}xy + \cdots + b_{ij}x^j y^i + \cdots.\end{aligned}\tag{3.7}$$

A similar representation was used by Gouesbet who started investigating under which conditions it was possible to reconstruct the vector field  $(X, Y, Z)$  from a single measurement  $X = x$  of an original vector field  $(x, y, z)$  [57, 58]. The reconstructed system was expressed as

$$\begin{aligned}\dot{X} &= Y, \\ \dot{Y} &= Z, \\ \dot{Z} &= F(X, Y, Z),\end{aligned}\tag{3.8}$$

and both polynomial and rational expansions for  $F$  were investigated. The polynomial expansions of  $F$  used were equivalent to (3.7). Although no reconstruction from data was performed in that paper, the challenge was acknowledged and pursued shortly after [59]. In the last paper, the important issue of estimating the derivatives based on finite-difference schemes from data was considered. This was an important step towards dealing with real data but a key hurdle still had to be transposed: measurement noise. To do this was one of the key objectives in [60], where this global modeling technique finally reached maturity. As pointed out in an early paper "noise removal ( $\cdots$ ) is required because standard systems ( $\cdots$ ) rely on derivative evaluations" [58, page 1795], to deal with noise was to *remove* it from the data, not so much as for discrete-time modeling techniques which are made to cope with noise.

Other polynomial expansions are to be found in the literature. Giona and colleagues investigated the use of a basis of polynomials to approximate both continuous-time and



**Figure 2:** Quotation diagram.  $A \rightarrow B$  indicates that A quotes B. The  $\star$  indicates that only paper [10] was quoted. Also, Crutchfield and McNamara quoted Farmer and Sidorowich's paper as the technical report: Preprint LA-UR-87-1502.

discrete-time dynamical systems [61]. In both cases the system was approximated by a linear combination of such polynomial functions which are obtained "from the knowledge of the hierarchy of moments." The procedure is quite involved, and long data sets are required to estimate the parameters. In their examples,  $N = 5 \times 10^4$  was used for the Hénon and Ikeda maps, and  $N = 4 \times 10^5$  was used for the Lorenz system.

Practical implementation of global modelling using continuous-time model requires an integration scheme. For instance, the explicit Euler integration scheme

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \mathbf{F}(\mathbf{y}_n) \delta t, \quad (3.9)$$

where  $\mathbf{y}_n = \mathbf{y}(n\delta t)$  with the time step may be used. It is known that the Euler is not very robust against time step change. This is why a Runge-Kutta scheme is most often used. Another alternative was proposed by [62]. They used an Adams-predictor-corrector scheme according to

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \delta t \sum_{j=0}^M a_j^{(M)} \mathbf{F}(\mathbf{y}_{n+1-j}), \quad (3.10)$$

where  $a_j^{(M)}$  are the implicit Adams predictor-corrector coefficients [62].  $M$  designates the order of the corrector portion of the integration. The global model  $\mathbf{F}$  has a polynomial form and is optimized with the help of a minimum description length criterion and the error function

$$\chi^2 = \frac{1}{2N\sigma^2} \sum_{n=1}^N \left[ \mathbf{y}_{n+1} - \mathbf{y}_n - \delta t \sum_{j=0}^M a_j^{(M)} \sum_{i=0}^{N_p} K_i \phi_i(\mathbf{y}_{n+1-j}) \right]^2 \quad (3.11)$$

which is quadratic with respect to coefficients  $K_i$ . This is thus a least square problem.

Structure selection issues for continuous-time polynomials have been discussed in [63, 64]. In [65] the authors use a priori knowledge of the driving force to help determining the model structure. When differential embeddings are used, the global model (3.8) is usually searched using a truncated polynomial expansion involving successive derivatives of the measured time series. In order to reduce the number of monomials, that is, the length of the model, it is possible to use a rational function whose monomials are selected according to an Ansatz library. Such a library contains canonical forms which can be mapped into an equivalent polynomial system whose variables are combinations of the measured time series and its successive derivatives. The canonical function  $F$  thus only contains selected terms. This reduced function allows to avoid numerical instabilities usually encountered with rational function. This procedure was introduced in [66] and improved in [67]. For instance a very accurate 7-term global model was obtained from the  $x$ -variable of the Lorenz system. A 26-term model was obtained from a copper electrodisolution experiments [68]. This model has to be compared to the 52-term model previously obtained in [69].

Other approximations for functions  $f$  and  $g$  in (3.7) include Volterra series expansions [70]. Examples of building global continuous-time models from real data can be found in a number of papers [69, 71–74].

### 3.2.2. Neural Networks

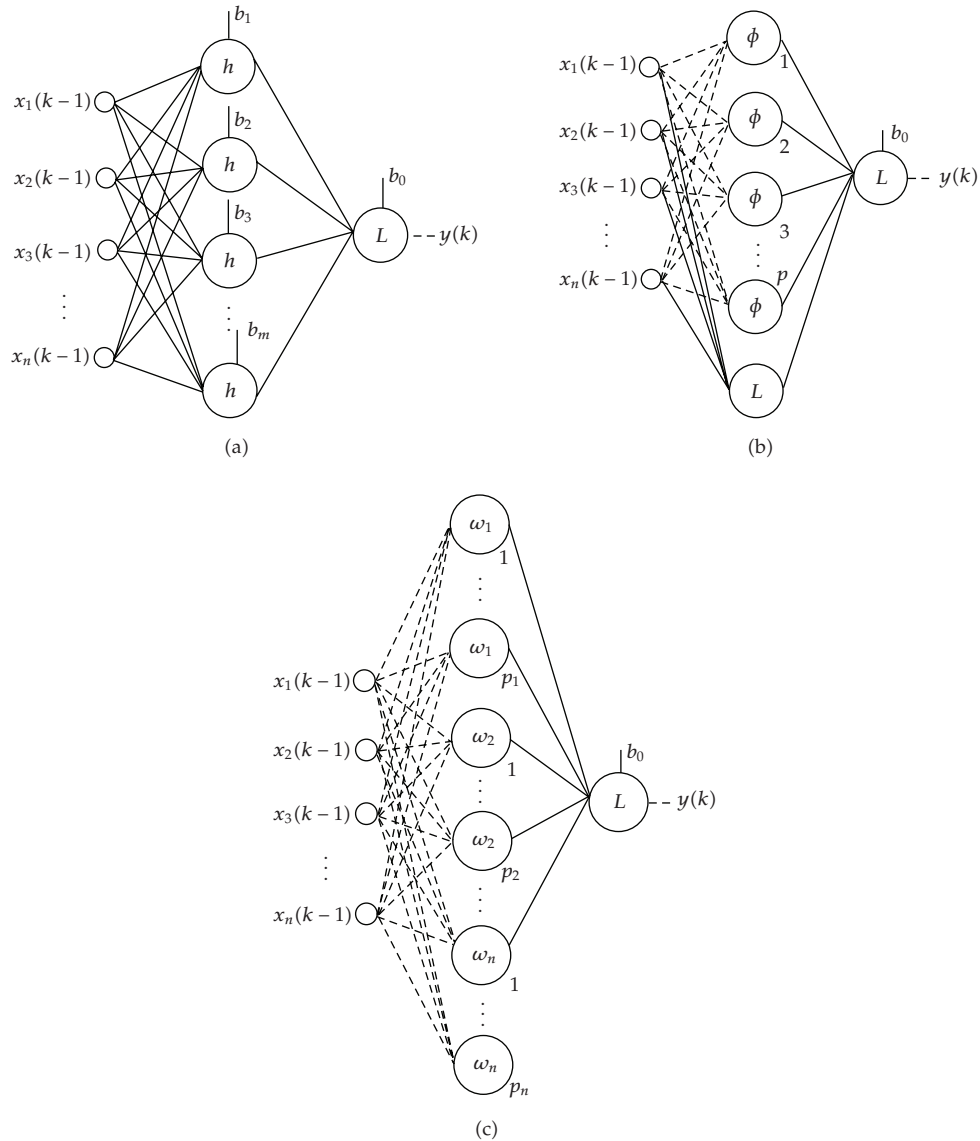
A neural network is a model class that resembles some aspects of a brain. Conventional simplifications made for perceptron models are: (i) to take only one hidden layer of nodes, (ii) to consider the output node linear, and (iii) to consider all the activation functions  $h$  of the hidden layer nonlinear are the same. A perceptron model with such features is illustrated in Figure 3(a) and can be described mathematically as

$$y(k) = b_o + \sum_{j=1}^m w_j^o h \left( b_j + \sum_{i=1}^n w_{ji}^h x_i(k-1) \right), \quad (3.12)$$

where  $w_{ji}^h$  indicates a weight (to be estimated) of the hidden layer that connects the  $i$ th input to the  $j$ th neuron of the hidden layer.  $w_j^o$  is the weight (to be estimated) of the  $j$ th hidden neuron output,  $b$ 's are constants, called bias parameters, and the neuron *activation function* is  $h$ . Finally,  $n = \dim(\mathbf{x})$  and  $m$  is the number of neurons in the hidden layer. The function shown in the right-hand side of (3.12) is often called *feed-forward* because there are no feedback loops *internal* to the network. It is important to notice that (3.12) is in the form  $y(k) = f[\mathbf{x}(k-1)]$ . Common choices for nonlinear activation functions are Gaussian, sigmoidal, and the hyperbolic tangent. In [75] they discuss the use of polynomials as activation functions for neural networks. An accessible introduction on the subject of neural networks applied to modeling chaotic systems is given in [76].

The first papers to build this kind of models from chaotic data seems to have been the references [78, 79]. (An earlier paper built and characterized a chaotic neural model with a single neuron but the network was not trained from data [80].) More recent and impressive results have been discussed in [81].

A rare study of bifurcation diagrams achieved by neural models has been presented in [82]. And an early work which reports on the improved performance of pruned networks is [83]. Neural networks have been applied in a number of instances, some examples include [84–86].



**Figure 3:** Schematic representation of some general model classes. (a) Typical perceptron model, (b) single-type basis function models, (c) multitype basis function models. The solid lines correspond to parameters that must be estimated. Dashed lines indicate absence of parameters. Therefore (a) is nonlinear in the parameters whilst (b) and (c) are linear in the parameters. Usually  $\Phi$  depends on some parameters. When this is the case, such parameters are chosen beforehand. Figure reproduced from [77].

### 3.2.3. Radial Basis Function Models

The RBF model class is shown in Figure 3(b). One important difference with respect to the first class is that now there are no weights associated to the connections between the inputs and the nodes in the hidden layer. The  $\phi$ —which is usually chosen as a radial function—is nonlinear and often depends on certain tuning parameters which are usually known when the weights associated to the connections indicated by solid lines (Figure 3(b)) are to be

estimated. As a consequence, the model is linear in the parameters and can be written thus (compare with (3.4) and (3.5))

$$y(k) = b_0 + \sum_i^p \omega_i \phi(\mathbf{x}(k-1)) + \sum_{i=1}^n a_i x_i(k-i), \quad (3.13)$$

where  $p$  is the number of radial basis functions.  $b_0$ ,  $\omega_i$ , and  $a_i$  are the unknown parameters to be estimated. Two of the first works to use this model class were mentioned in Sections 3.1.3 and 3.1.4 above. Work using this type of model has been recently surveyed in [87].

A number of papers have described the application of RBF networks in the modeling of nonlinear dynamical systems and chaos [88–91]. An accessible introduction on the subject of RBF models applied to modeling chaotic systems is given in [76].

### 3.2.4. Discrete-Time Polynomials

The model class illustrated in Figure 3(c) is quite similar to the preceding one. The main difference is that the basis functions are usually different, that is,  $\omega_i \neq \omega_j$ . Another important difference, not revealed in the figure, is that, whereas it is usually assumed in the second class that the input vector is uniform, in the third class such uniformity is not required (see Section 4.1). However, the main difference is that various basis functions are often used in Figure 3(c) in such a way as to enable matching different data features. One possible choice of basis functions  $\omega_i$  is monomials of different degrees of nonlinearity in the range  $1 \leq m \leq \ell$ . Each  $m$ th-order term can contain a  $p$ th-order factor in  $y(k-n_i)$  and a  $(m-p)$ th-order factor in  $u(k-n_i)$  and is multiplied by a coefficient  $c_{p,m-p}(n_1, \dots, n_m)$  as follows:

$$y(k) = \sum_{m=0}^{\ell} \sum_{p=0}^m \sum_{n_1, n_m}^{n_y, n_u} c_{p,m-p}(n_1, \dots, n_m) \prod_{i=1}^p y(k-n_i) \prod_{i=p+1}^m u(k-n_i), \quad (3.14)$$

where

$$\sum_{n_1, n_m}^{n_y, n_u} \equiv \sum_{n_1=1}^{n_y} \cdots \sum_{n_m=1}^{n_u}, \quad (3.15)$$

and the upper limit is  $n_y$  if the summation refers to factors in  $y(k-n_i)$  or  $n_u$  for factors in  $u(k-n_i)$ . In the case of noisy data  $y(k-n_i)$ , noise terms *must* be included in (3.14) to avoid the error-in-the-variables problem (see Section 5).

The choice of monomials as basis functions constrains the resulting models to those cases in which the dynamics underlying the data can be approximated by a linear combination of nonlinear monomials. For systems that are more strongly nonlinear, other basis functions should be preferred. On the other hand, the choice of monomial basis functions enables building models which are more information dependent than models for which all the basis functions are of the same type.

Discrete-time polynomial models have been used in a number of papers within the field of nonlinear dynamics and chaos. Bagarinao and colleagues used this model class to reconstruct bifurcation diagrams from data [92–95]. Data analysis and modeling applications

include [96–103]. Applications to real data—in the field of nonlinear dynamics—include [104–106].

### 3.2.5. Rational Models

Rational models are composed by the division of two polynomials such as (3.7) in the continuous time, and such as (3.14) in discrete time. It should be noticed at once that whereas *deterministic* polynomial models are linear-in-the-parameters, rational models, even deterministic, are nonlinear-in-the parameters. This calls for a series of cares in their estimation [107–109].

Rational models for continuous-time systems were considered in [57], although only analytically, as no model building was attempted in that paper. Rational models were estimated from data produced by an electronic oscillator in [110]. The practical difficulties with rational models seem to be revealed by the small number of papers that deal with such models in the context of nonlinear dynamics and chaos. At least for continuous-time model, rational functions lead to many numerical instabilities that are not trivial to remove. Nevertheless, such a problem is no longer observed when return maps are considered. Thus, Lorenz map and Ikeda map have been successfully reproduced with rational models [111].

### 3.2.6. Wavelet-Based Models

Wavelet models, wavelet networks, or just *wavenets* are similar to the linear-in-the-parameter models described in Sections 3.2.3 and 3.2.4. The main difference is that instead of radial basis functions or monomials, wavenets are composed of a linear combination of a set of wavelet functions  $\psi_{a,b}(t)$  formed from a so-called wavelet prototype  $\psi(t)$  by dilations and translations such as

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), \quad (3.16)$$

where  $t, a, b \in \mathbb{R}$  and  $a > 0$ . As for RBF models, there are various choices of the wavelet function  $\psi(t)$ . There seems to be no clear guidance as to which type of function to choose. Since the various functions  $\psi_{a,b}(t)$  used to compose the model are different (for different  $a$  and  $b$ ), then the structure selection for this type of models resembles that of the discrete polynomials described in Section 3.2.4 because the user must chose not only the number of basis functions to use, but also the features of them (the  $as$  and  $bs$ ).

One of the first papers to use wavenets in order to model a chaotic system was [29]. The authors used a mexican hat type wavelet function  $\psi(t)$  to model various benchmark models. Optimized tensor product wavelet models were obtained for data from a vibrating string experiment in [112], and B-splines wavelet models for the nonautonomous Duffing oscillator were discussed in [113]. Wei and Billings investigated the use of structure selection algorithms for wavelet models in the context of chaotic systems [114].

### 3.2.7. Fuzzy Logic

Fuzzy models use internal variables which are linguistic. At a certain point of modeling the numerical variables must become linguistic by taking labels such as small negative large positive. The core of a fuzzy model then consists of a set of rules of the type: if  $x$  is large, then  $y$  is medium. Finally, the linguistic variables must be changed back to numerical variables.

Of course there are a number of subtleties which actually make this type of modeling work [115–118].

From the list of model classes surveyed in this paper, it seems to us that fuzzy logic models have been the least used to model nonlinear dynamics and chaos. One of the first papers on the subject seems to be [119]. A comparison of several techniques, including fuzzy models, in a problem of nonlinear prediction can be found in [120].

### 3.2.8. Input-Output Models

Input-output models are not a class of models but simply mean that the model class caters for the use of input signals. By input we mean external, time-dependent signal(s). Therefore an input-output model will be nonautonomous.

For the modeling of input-output models it is necessary to record not only the output, but also the input. As a general rule, global differential equations built from data are autonomous (for an exception, see [65] where the method applies to harmonically driven systems). In [121] the authors discuss the use of a bifurcation parameter (as an input) in the estimation of differential equations. However, in general, if inputs must be handled, then discrete-time models are more convenient.

There have been attempts to develop an embedding theory for input-output models [122, 123]. The feasibility of this theory and its benefits are not totally clear. Fortunately, there are other theoretical approaches to the input-output case [25, 124], and a vast number of examples illustrate the practical value of such models.

Nonautonomous chaotic models have been obtained in [65, 113, 125].

### 3.2.9. Equivalence between Continuous and Discrete-Time Polynomials

As surveyed above, there are various methods for building continuous-time models, typically in the form of ODEs. However, the data available is always discrete in time, and the simulation of ODEs is also carried out on digital computers. Therefore, at some stage some type of discretization of the ODEs must occur. In some methods such discretization is intentional [62, 126], however, in such cases the final aim was to have an ODE at the end of the day. What happens to the ODEs fixed-points after it is discretized was discussed in [127].

A more fundamental question concerns the choice of the integration step of numerical integration schemes. Given an ODE, say that produces a chaotic attractor, what happens to the attractor if the integration is varied? If it becomes too great, certainly the attractor will change, but will the “new” attractor still be some attractor of the original system? These questions have been addressed in [128]. In fact, so long as the frequency corresponding to the integration step is more than twice the highest frequency in the Fourier spectrum (Nyquist criterion), the “new” attractor is still topologically equivalent to an attractor solution to the set of ODEs. Only a displacement in the parameter space is induced by the integration step, which functions as a bifurcation parameter in some cases. It is only when the integration step is larger than the maximum established by Nyquist’s criterion that the obtained attractor is spurious, that is, associated with numerical instabilities.

## 4. Structure Selection

In few words, the problem of structure selection is that of deciding *how many* and in some cases *which* function basis (see, e.g., Figure 3(c)) should be used to build a dynamical model

from data. This problem gains different tones depending on the model class considered but it is always present. Often in the context of networks, structure selection is known as “topology determination.” This issue has been addressed in a number of papers [129–133]. In the case of fuzzy models, structure selection boils down to define how to partition the input space and to define the number of rules. Structure selection issues for fuzzy models have been investigated in [35, 134, 135]. Even for local linear models, structure selection would be the definition of the embedding and the number, size and location of neighborhoods, and so forth.

In the field of system identification, the issue of structure selection for nonlinear models gained much attention by the late 80 s. However, in the field of nonlinear dynamics and chaos this trend was delayed a few years. In the early (and sometimes late!) 90 s several papers simply assumed the structure known [55, 136–138]. A good example of what happens when structure selection is not considered during model building is provided in [139].

Soon the practical importance of model structure selection started to be recognized and dealt with in various ways and in varying degrees of success [62, 140, 141]. By the mid 90 s most authors became aware of the necessity of structure selection regardless of the model class chosen. However, the problems due to overparametrization were not well diagnosed, usually being attributed to an allegedly poor model class, poor data, and so on. A detailed study of the dynamical effects of overparametrization on model attractors and bifurcation diagrams was produced in [142]. Other similar studies followed [143–147].

The key point in structure selection is to choose a model structure that is as simple as possible, but also sufficiently complex to capture the dynamics underlying the data. One of the easiest (and less efficient) methods for model selection is called zeroing-and-refitting [141] which consists of estimating the parameters for a large model—which is already a problem—and to eliminate those which are “sufficiently small.” Of course, the “size” of a parameter will depend on the particular window of data used, on the noise variance, and on the particular type of basis function used, especially if the data are not normalized [148]. Therefore, zeroing-and-refitting does not generally work in practice [62].

An alternative and simple way of tackling this problem for nonlinear systems has been proposed in [149]. Instead of deleting specific model terms, entire term clusters should be deleted, based on the behavior of the respective cluster parameters. Also, certain term clusters can be deleted to force symmetry [144, 150], which is necessary in some cases [151]. The concept of the nearest neighbors was used in [152] to determine the best input and output lags in NARX (nonlinear autoregressive with exogenous inputs) models.

One way of addressing the structure selection problem are to define some measure of complexity for a given model. In their paper Crutchfield and McNamara were concerned with quantifying and limiting the complexity of their models. They chose models that minimized the model entropy  $I(M)$  and argue the importance of such a measure in the context of chaotic data [37]. The *maximum description length* (MDL) [153] has been used in several papers [132, 154]. Other ways of tackling the structure selection problem is to define a measure of quality for each regressor *in an orthogonal space* and then use such a criterion to select those regressors that are most relevant. This is the basic idea behind the *Error Reduction Ratio* (ERR) proposed in [155] and used in [156].

Unfortunately, there is no definite solution to the model structure selection problem so far. Situations in which the current methods fail abound. Brown and colleagues report failure of the MDL criterion [150], and Piroddi provides simple examples in which the ERR fails [33]. Fortunately, the benefits and successes outnumber the failures. New algorithms for structure selection are published at an amazing pace [30–35]. The authors of this survey are convinced



that the use of a priori information (see, e.g., [144, 151, 157]) will prove useful in addition to the new techniques.

#### **4.1. Uniform and Nonuniform Embeddings**

A key issue in modeling nonlinear dynamics is that of selecting an appropriate embedding space. In principle, this would include two stages: the choice of observables [158, 159] and the choice of embedding parameters [160]. In many practical situations, although the observable is determined before data acquisition, the embedding parameters—basically the embedding dimension and the delay time—can be determined by the user a posteriori. In [161] a method has been suggested to try to detect if a given variable is appropriate for global modeling. Using the nomenclature of Figure 3, the input vector, at the left-hand side of the models, determines the embedding space.

It has been duly pointed out that the problem of choosing an embedding in the context of model building is a bona fide stage of the modeling procedure [162]. Also, the uniform embedding defined by taking the elements of  $\mathbf{y}(k)$  in (3.1) to be the coordinates is not necessarily optimal. Therefore, which elements (coordinates) should be chosen to compose  $\mathbf{y}(k)$  is also part of the modeling problem. An optimal solution to such a problem might require an embedding space in which the “temporal distances” from one coordinate to another are not necessarily the same. The authors of [162] classify such embedding spaces as irregular. Despite this, to build discrete-time models using regular embeddings seems to be the rule rather than the exception in most of the literature.

To enable irregular embeddings (see Figure 3) it suffices not to connect certain input nodes to the middle layer. Here it is pointed out that the choice of particular basis functions  $\omega_i$  is, in some cases equivalent to the choice of embedding coordinates which could turn out to be irregular. This is one of the advantages of using the ERR criterion to accomplish structure selection. In fact, (13) and (21) of [156] and (37) of [125] are some examples of models (automatically) built on irregular embeddings. In conclusion, it becomes clear that the modeling procedure followed in [125], for instance, includes the choice of embedding coordinates as a part of the modeling procedure, as pointed out in [162].

### **5. Parameter Estimation**

Before, it is noted that synchronization has been used in parameter estimation problems [163–166].

One of the most commonly used algorithms for estimating unknowns in linear-in-the-parameters models is the least-squares algorithm or some generalization of it [34]. In the case of noisy measured data the least-squares estimator is biased due to the error-in-the-variables problem [167, 168] because such an estimator does not take into account the measurement errors [169]. Fortunately, there are well-established and robust algorithms for such situations which are mildly nonlinear [22]. Some of such algorithms solve the optimization problem in a higher-dimensional space thus successfully avoiding bias. Due to the unbiased estimators—some of which take into account the measurement errors—in the field of system identification, the error-in-the-variables problem only occurs when both input and output are noise contaminated. When only the output is noisy, the unbiased estimators successfully circumvent the error-in-the-variables problem.

All such algorithms are based in some norm of one-step-ahead prediction error. An alternative would be to minimize some norm of a  $k$ -step-ahead prediction error, with the

advantage of gaining robustness to noise. In such a case, however, the resulting optimization problem becomes strongly nonlinear and should be solved with adequate tools. The use of a back-propagation algorithm in this context was used in [170]. In [43], which also followed a local model procedure, the local maps were fitted to the data in such a way as to be consistent with  $k$ -step-ahead predictions. The authors of that paper reported that estimating parameters by least squares did not always yield good results. This should come as no surprise because they were fitting local maps and therefore required additional information to constrain parameter estimation. Global models are less sensitive to this type of problem.

In what concerns parameter estimation, an important difference between [59, 60] and [56] is that, whereas the former uses linear estimators—at the expense of having to estimate derivatives from data—the latter numerically integrates (3.7) and uses the observed data as they are. This procedure, which was called the *trajectory method*, results in an optimization problem which the authors solve using standard routines in the IMSL10 library. In [55], the authors provide more information on this estimation procedure and also consider the discrete-time case. The trajectory method in the context of RBF models is discussed in [76]. It should be noticed that [55, 56] do not consider modeling problems such as structure selection. The application of the trajectory method was illustrated in [171], and the use of the multiple shooting approach to estimated parameters of ordinary differential equations was discussed in [73, 136, 169].

For *neural networks* the weights and the bias terms are determined by optimization algorithms that search to minimize a cost function which usually depends on the difference between the given data and the network output. Because such models are nonlinear in the parameters, the optimization problem must be solved using some kind of nonlinear estimator, such as the back-propagation algorithm. In a recent study, different approaches to network training were compared [172].

Breeden and Packard have discussed the use of genetic algorithm and evolutionary programming for solving a number of optimization problems which occur in the modeling of nonlinear dynamics and chaos [173]. Other parameter estimation techniques include a generalized Gauss-Newton method for maximizing a likelihood function [74]; ridge regression or regularization [49, 89]. The multiple shooting approach has been used and discussed in a number of papers [136, 174]. Nonparametric estimation (a rare example in nonlinear dynamics) was provided in [175]. The use of regularization has been considered in [49, 76]. A specific procedure for estimating three parameters of a differential-delay equation was put forward in [176] and another method, specific to a certain class of one-dimensional maps, was put forward in [177]. A multiobjective estimator was discussed in [178]. Finally, a number of issues directly concerned with the estimation of dynamical invariants and indirectly related to the estimation of model parameters were discussed in [167].

## 6. Model Validation

The issue of model validation is vast. In order to cover such a wide subject in limited space, we will base this section on the paper [179] and refer the reader to [180] for a coverage of the field up to the beginning of the nineties.

Before actually starting to describe some results in the literature, a few remarks are in order. First and foremost, the challenge of model validation or of choosing among candidate models should take into account the intended use of the model. Hence, a model could be good for one type of applications and, nonetheless, perform poorly in another. In the context of this paper, the main concern is to assess the model dynamics. A different concern, though equally

valid, that would probably require a different approach would be to assess the forecasting capabilities of a model. Second, it should be realized that two similar though different problems are: (i) model validation, which usually aims at an *absolute* answer like: valid or not valid, (ii) model selection, which usually aims at a *relative* answer such as: model  $\mathcal{M}_1$  is better than model  $\mathcal{M}_2$  according to criterion  $C$ . Finally, when it comes to model validation, the safest approach is to use many criteria, rather than just one.

Although very popular in other fields, the computation of various *measures of prediction errors* (one-step-ahead and free-run) in the case of nonlinear dynamical systems is not conclusive in what concerns the overall dynamics of the identified model [81, 132, 180], though it does convey much information on the forecasting capabilities of a model.

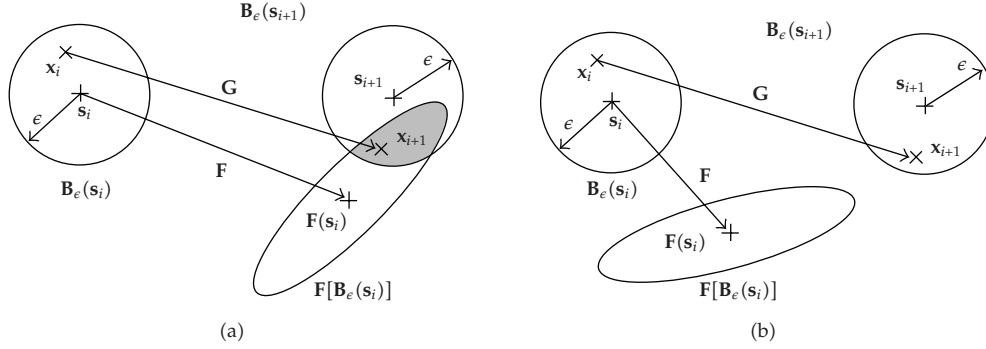
Subjective though it is, the *visual inspection of attractors* (or simply comparing the morphology of two time series) is quite common a way of assessing the quality of models [52, 60, 65, 72, 73, 78, 91, 119, 132, 162, 168, 171, 174, 181–183]. Such a procedure is not only subjective but also ineffective to discriminate between “close” models, that is, models with slight, but important, differences in their dynamical behavior. What renders this procedure subjective is the fact that no quantitative mechanism is used to compare how close are two reconstructed attractors. In this respect the work by Pecora and coworkers could be an alternative for determining how close are the original and the model attractors [184]. To the best of our knowledge the statistic measures that put forward in the mentioned paper have not yet been used in the context of model validation.

Still in relation to the visual inspection of attractors, it should be noticed that in many practical instances there is not much more that can be done consistently. For instance, in the case of slightly nonstationary data, to compare short-term predictions with the original data is basically the best that can be done. Building a model for which the free-run simulation approximates the original data in some sense is usually a nontrivial achievement. (In this respect we rather disagree with [185] who consider free-run simulation of models a trivial validity test.)

Other *attractor features* are still in common use when it comes to model validation. Among such features the following are frequently used: (some of such properties have been recently discussed in the context of model validation in [186]) Lyapunov exponents [43, 81, 97, 145, 147, 187, 188]; correlation dimension [78, 81, 132, 188]; location and stability of fixed-points [112, 125, 183]; (In particular, it has been shown that fixed-point stability of nonlinear models is consistent with breathing patterns found in real data [189].) Poincaré sections [81, 176]; geometry of attractors [85]; attractor symmetry [150, 151]; first-return maps on a Poincaré section [66, 190]; probability density functions of recurrence in state-space [191]; topological features such as linking numbers and unstable periodic orbits (UPOs) [192–194].

Meaningful validation can only be accomplished by taking into account the intended use of the model. A model that provides predictions consistent with the observed data will probably not be a good model to study, say, the sequence of bifurcations of the original system.

The use of model free-run simulations in the context of surrogate data analysis for model validation was suggested in [195]. For details on surrogate data analysis the reader is referred to [196, 197] and for a chemical process application; see [198]. The main idea is to use estimated models to produce a large number of time series and to use some test statistic to try to assess if it is likely that the data could have been produced by a model (or models) such as those used to produce the surrogates. Of course, a key point in this procedure is the choice of the test statistic. For instance, suppose that the correlation dimension or the largest



**Figure 4:** Graphical interpretation of (a) consistent prediction and (b) inconsistent prediction [200].  $\mathbf{G}$  and  $\mathbf{F}$  are the “true” dynamics and model, respectively.  $x_i$  is the true state at time  $i$  and the observed state at that time is  $s_i$ .  $\mathbf{B}_\epsilon(s_i)$  is an observational uncertainty sphere of radius  $\epsilon$ . In (a) the shaded region indicates that there is a subset of forecasts made with  $\mathbf{F}$  that are indistinguishable from the “true” state ( $x_{i+1}$ ) within observational uncertainty. Figure reproduced from [179].

Lyapunov exponent is chosen to compare a set of free-run simulated data with the measured data. Suppose further that the null hypothesis is that the measured data is compatible with the estimated model. Even if we cannot reject the null hypothesis, that does not guarantee equivalence of *dynamical* behavior because quite different attractors may have similar indices [199]. This is also true for the correlation dimension.

A related method has been described in [200] and has been named consistent nonlinear dynamic (CND) testing. The main idea is depicted in Figure 4.

In the following two subsections we briefly discuss two other procedures which the authors have developed for model validation. The first one, following an early paper by Brown and colleagues [201], is based on the concept of synchronization. The second procedure, which is much more demanding, but also provide a much deeper insight into the model dynamics is based on topological analysis.

### 6.1. Synchronization

As done in Figure 4, let us denote the “true” dynamics by  $\mathbf{G}$  and a given model by  $\mathbf{F}$ . In the present section it is assumed that the dynamics are continuous, that is [201],

$$\begin{aligned} \frac{dx}{dt} &= \mathbf{G}(x), \\ \frac{dy}{dt} &= \mathbf{F}(y) - \mathbf{E}(y - x), \end{aligned} \quad (6.1)$$

where it has been assumed that  $\dim[x] = \dim[y]$  and where the matrix  $\mathbf{E}$  denotes the coupling between the true system and the model. The scheme illustrated in (6.1) will be referred to as dissipative synchronization or entrainment.

The rationale behind this procedure is as follows. Assume that the data  $x$  lies on a chaotic attractor. In many situations, provided  $\mathbf{E}$  is adequately chosen and  $\mathbf{G} = \mathbf{F}$ ,  $y \rightarrow x$ . That is, the model will synchronize to the system. If  $\mathbf{G}$  and  $\mathbf{F}$  differ slightly, the error  $e = y - x$  will not go to zero but will stay around the origin of the error space. The average distance to the origin of such a space will depend on  $\mathbf{G}(x) - \mathbf{F}(x)$ . Therefore such a distance (which in

practice is a measure of quality of “synchronization”) is a measure of how far the estimated model  $\mathbf{F}$  is from the true dynamics  $\mathbf{G}$  [201].

The difference  $|\mathbf{x} - \mathbf{z}|$ —where  $\mathbf{z}$  is the model state vector *without* any driving force—is compared to  $|\mathbf{x} - \mathbf{y}|$ . If  $|\mathbf{x} - \mathbf{y}|$  drops below a certain threshold ( $10^{-2}$  was used in [185]) and is “clearly” smaller than  $|\mathbf{x} - \mathbf{z}|$ , that is,  $\mathbf{x} \approx \mathbf{y}$  and it is assumed that the model is synchronized to the data, therefore  $\mathbf{F}$  should be sufficiently close to  $\mathbf{G}$ .

As it often happens in the realm of model validation, this procedure also is highly subjective, since it requires an ad hoc threshold, mentioned in the previous paragraph. In what concerns dissipative synchronization, it is well-known that in many cases by increasing the strength of the coupling (matrix  $\mathbf{E}$ ) it is possible to force a greater degree of synchronization and, in some cases, even attain identical synchronization [201]. For instance, in [185] the authors found that for values of the coupling greater than 2, models of an electronic circuit would synchronize with the measured data. On the other hand, in [202] they have found a lower bound of 0.1 for the coupling strength in order to guarantee synchronization between the Rössler system and perturbed versions of the original equations. It therefore becomes clear that it is sometimes possible to synchronize even a poor model to the data so long as the coupling strength is made sufficiently large. In fact, it has been shown that even different systems can synchronize, at a rather high cost [203].

Therefore although the concept of synchronization could be useful in the context of model validation, it becomes apparent that some adjustments are required to render the procedure more practical. In [179] some steps were given in this direction, and it was shown using numerical examples that synchronization yielded similar results to the use of bifurcation diagrams for model validation [180] but at a much lower computational cost. Bifurcation diagrams in model validation were used in [86, 89].

## 6.2. Topological Analysis

A topological analysis usually starts by computing a first-return map to a Poincaré section of the phase portrait. This is indeed useful for extracting periodic orbits using a close return method although that other techniques can be used. In the best cases, the first-return map is made of  $n$  monotonic branches separated by an  $n - 1$  critical point. The first-return map induces therefore a partition of the phase portrait in  $n$  zones. A symbol is associated with each branch.

Chaotic trajectories and the periodic orbits constituting their skeleton are thus encoded over the symbol set  $\{0, 1, \dots, n - 1\}$ . Even symbols are associated with increasing branches. Increasing branches are preserving order and decreasing branches are reversing order. The phase portrait is thus divided in preserving order and reversing order strips. A preserving order strip presents an even number of half-turns while a reversing order strip represents an odd number of half-turns. The attractor can thus be schemed by a branched manifold—a template or a knot holder—on which periodic orbits can be drawn.

All topological properties are encoded in the template. Thus it is possible to extract topological invariants like linking numbers from a template construction. Linking number between two periodic orbits is the most often used topological invariant. It can be counted on a regular plane projection of the two periodic orbits. Each crossing (in the plane projection) between the two orbits is associated to  $\pm 1$  according to the third coordinate of each orbit segment. The linking number is then the half sum of the orientated crossings. The template is valid when all linking numbers it predicts are equal to those counted from the periodic orbits extracted from the attractor studied. Many details about topological analysis can be found in [204].

Topological analysis is surely the strongest validation method. Unfortunately, it is up-to-now restricted to 3D dynamics. Topological validation of global models was performed for the copper electrodisolution experiments [69], a string experiment [193], a Belousov-Zhabotinskii reaction [205], and so forth.

## 7. Modeling with A Priori Knowledge

We start this section with a quotation from one of the pioneering papers of the field: “If extra information is available about a system in addition to a time series, such as explicit underlying partial differential equations or symmetries, then the inverse approach as presented here does not exploit such information. Consequently, for such systems it may be possible to improve significantly upon the inverse approach using other techniques” [39, page 354].

Despite opinions as the one just quoted, the problem of building models from data *and* additional information (sometimes referred to as a priori or *auxiliary*) has been postponed. In the context of linear models or nonlinear process models, some results are available [206–210]. However, as pointed out in [211] there seems to be less applications of gray-box modeling in the realm of nonlinear dynamics.

The general setting is to have the dynamical data *plus* some other source of information and to use both in building a model. The knowledge of the model equations is optimistic to be considered a priori information, and here it is assumed that structure selection techniques (Section 4) will be used to find out which basis functions should be used to compose the model. If all that is desired of the system is already available in the particular set of data used for model building, then granted that the model class is sufficiently general and that the model structure is adequate, the resulting model should be a sufficiently accurate representation of the system. In this case there is no motivation to use an additional piece of information. However, often in practice, for a number of different reasons—such as presence of noise, poor choice of observable, poor frequency content in the data, limited amplitude excursion, and the like—the available data either does not have all the desired information about the system or such information is difficult to obtain. In such cases it is conceivable that additional information is available and it is natural to enquire if it is possible to build the model. Thus, more often than not, if the data is of “good” quality and “sufficiently” complete, black-box modeling should be the practitioner’s first choice.

In the realm of nonlinear dynamics, procedures have been put forward for building models using auxiliary information. A number of fixed-points were used in [144], the location of fixed points were used in [178]. Information about the harmonic driving were used to choose the model structure in [65]. Information about the symmetry was used to constrain not only the topology but also the parameter estimates of network and radial-basis function models [151]. In the last quoted paper, for instance, symmetry was imposed on a multilayer perceptron neural network in order to guarantee the pitchfork bifurcation—which is structurally unstable—when modeling the Duffing-Ueda oscillator. Topological features such as folding and tearing mechanisms [199, 212] in addition to the location and local eigenstructure of fixed points have been used in [213]. Lastly, information about the first period-doubling bifurcation diagram was used in [77].

## 8. Remaining Challenges and Conclusions

As early as 1987, the issue of modeling spatiotemporal nonlinear dynamical systems can be found in the literature [37]. Defining which type of a priori information is usable and then

learning how to use it for a given model class remain a challenge. In this difficult problem it is necessary to find adequate duplets of information and model class. To possess a priori information and not knowing how to use it for a given model class is useless. The reverse is equally true. The combination of models with different properties in an assemble approach [52] seems a promising line for future research in nonlinear dynamics and chaotic systems.

Over fifteen years ago Robert Gilmore, referring to model building, mentioned: "There are at present two distinct methods to model data. One is analytic, and has not been extensively used. The other is topological, and has been used even less" [214, page 515]. Although this paper did not aim at describing in detail the significant progress made in the field of "topological analysis," hopefully the reader did get the feeling that much has been done. On the other hand, concerning the "analytic methods" for modeling data, which was the main focus of this paper, it should become clear that quite a lot has been achieved since Gilmore's summary was written.

The view presented in this work, as any review paper, is strongly influenced by the authors' experience. A varied list of introductory and review articles was included. A rather short list of benchmark models and widely available data with indication of works that have used such models and data were provided. Although every effort has been made to survey such a wide subject in a comprehensive way, it goes without saying that a complete list of papers is outside the authors' reach. Nevertheless we are sure that the present survey will serve as a good starting point for future works in modeling nonlinear dynamics and chaos.

## **Appendix**

### **A. Benchmark Models and Data Sets**

As a helpful aid to those involved in developing new tools for modeling nonlinear dynamics and chaos, in this appendix we list some benchmark toy models and benchmark data sets (widely available) and point out papers that have discussed their modeling and analysis.

#### **A.1. Maps**

##### *Logistic*

The logistic map was originally investigated by Robert May [215, 216]. Papers that investigated the estimation of models from data produced by the logistic map include [38, 125].

##### *Hénon*

The logistic map was originally proposed in [217]. It was considered as a benchmark for modeling in [61, 125, 168, 218]. This map was modeled using a kernel estimation approach in [43], recurrent networks in [84], and a wavelet model in [114].

##### *Ikeda*

The Ikeda map was introduced in [219] and considered as a benchmark in [39, 61]. In [29] this map was modeled using wavelet networks, and in [132] results are reported for neural networks.

## **A.2. Autonomous Systems**

### *van der Pol's Oscillator*

The autonomous van der Pol oscillator settles to a limit cycle and therefore can be modeled. This oscillator has been considered in [119].

### *Lorenz*

The Lorenz system was introduced in [4] and considered as a benchmark in [39, 76], although only model prediction errors are given. The polynomial and rational expansions of that system from a single observable were considered in [58], and the construction of ODEs from data was presented in [59, 60, 62]. This system was modeled using a kernel estimation approach in [43], in [125] it was modeled using discrete-time polynomials, and in [29] it was modeled using wavelet networks.

Fuzzy models for this system were built in [119]. A piece-wise affine model for this system was developed and analyzed in [213], in a sense such a model bears resemblance with Takagi-Sugeno fuzzy models [116]. Unconstrained [85] neural networks and constrained RBF models [151] were also used in the modeling of this system.

### *Rössler*

The well-known Rössler system was proposed in [220]. The polynomial and rational expansions of that system from a single observable were considered in [57], and the construction of ODEs from data was presented in [59, 60]. Fuzzy models for this system were built in [119, 125] it was modeled using discrete-time polynomials. A piece-wise affine model for this system was developed and analyzed in [213], and a neural network was reported in [132].

### *Makey-Glass*

The Makey-Glass delay-differential equation was introduced in [221] and considered as a benchmark in [39, 76, 125]. In [29] this system was modeled using wavelet networks, and in [84] the reader will find results with feedforward and recurrent neural networks.

## **A.3. Nonautonomous Systems**

It must be pointed out that a nonautonomous oscillator of order  $n$  driven by a single frequency, say, an input of the type  $\cos(\omega t)$  is, in practice, an autonomous system with dynamical order  $n + 2$  [191]. In such cases the resulting models do not have an external variable  $u(t)$  or  $u(k)$  which can be *any* time-dependent signal.

### *Duffing's Oscillator*

This oscillator was considered for a fixed input of the type  $\cos(\omega t)$  in [37]. In [125] a model valid for any input was built from data. The estimated model and original oscillator have very similar bifurcation diagrams. Wavelet models with similar features were obtained in [113]. A close look at the pitchfork bifurcations will reveal that they are not (the symmetrical)



pitchforks but rather (the unsymmetrical) saddle-node bifurcations. Because symmetry is structurally unstable, any small deviation due to noise or “imperfect” estimation from data will destroy it, and with it the hope to reproduce a pitchfork bifurcation. However, by imposing symmetry it is possible to recover the correct bifurcation pattern. This has been done with neural network models for the Duffing oscillator in [151].

#### *van der Pol's Oscillator*

This oscillator appeared in [222]. An experimental implementation of this oscillator was considered for a fixed input of the type  $\cos(\omega t)$  in [37]. In [125] a model valid for any input was built from data. The estimated model and original oscillator have very similar bifurcation diagrams. Wavelet models for this oscillator have been obtained in [114].

### **A.4. Benchmark Measured Time Series**

Several papers in the literature illustrate the proposed techniques using measured time series. It would be impossible to list all of them. In this subsection, however, we list those time series that are most commonly used because they are widely available and therefore constitute typical benchmarks.

#### *Santa Fé Time Series Competition Data*

In 1994 the Santa Fé Institute promoted a time series prediction competition, which is described in [223]. Some measured and simulated time series were made available as benchmarks for the competition. The description of these data and the files can still be found at <http://www-psychem.stanford.edu/andreas/Time-Series/SantaFe.html>.

The Laser data was considered in [81, 91]. The biomedical data set has been described in detail by Rigney and coworkers [224] in view of the Santa Fé Time Series Prediction and Analysis Competition [225]. Such data were considered in [189, 226–230].

#### *Wolf's Sunspot Number Time Series*

These data can be found at <http://www.ngdc.noaa.gov/> and have been considered in [20, 49, 91, 106, 132, 231, 232]. In the last two references, the original data was transformed to a symmetrical space in which the modeling is easier. The transformed data can be found at <http://www.atomosyd.net/>.

#### *Copper Electrodeposition Data*

These data can be found at <http://www.atomosyd.net/spip.php?article40> and have been considered in [68, 69]. They were recorded by Zihao Fei and Jack Hudson. A 52-term global continuous-time model [69] and a 26-term model [68] were obtained from the data. A structure selection using an Ansatz library was used in the second case.

#### *Canadian Lynx*

These data results from the records of the Hudson Bay Company regarding the populations of lynx and hares. They can be found at <http://www.atomosyd.net/spip.php?article39> and have been considered in [233]. They were recorded by Zihao Fei and Jack Hudson.

### Electronic Oscillators

The electronic oscillator described in [234] was modeled and analyzed in [71, 73]. These data can be found at <http://www.y2k.maths.ox.ac.uk/>.

The so-called Chua's circuit [235] was modeled and analyzed in [236] using discrete-time polynomials, in [64] using continuous-time polynomials and in [110] using rational models. Neural networks from simulated data of this circuit are described in [237]. These data can be found at <http://www.cpdee.ufmg.br/~MACSIN/>.

### Acknowledgments

This work has been partially supported by CNPq and CAPES (Brazil) and CNRS (France). The authors are grateful to the editors for their encouragement and continued assistance in the preparation of this survey.

### References

- [1] H. Poincaré, "Sur les courbes définies par une équation différentielle," *Journal de Mathématiques Pures et Appliquées*, vol. 4, no. 1, pp. 167–244, 1885.
- [2] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste*, vol. 3, Gauthier-Villard, Paris, France, 1899.
- [3] A. A. Andronov, A. A. Vitt, and S. E. Khaikin, *Theory of Oscillators*, Pergamon Press, Oxford, UK, 1966, Traduction anglaise par F. Immirzi, Dover, New York, NY, USA, 1937.
- [4] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.
- [5] D. Ruelle and F. Takens, "On the nature of turbulence," *Communications in Mathematical Physics*, vol. 20, pp. 167–192, 1971.
- [6] R. M. May, "Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos," *Science*, vol. 186, no. 4164, pp. 645–647, 1974.
- [7] O. E. Rössler, "An equation for continuous chaos," *Physics Letters A*, vol. 57, no. 5, pp. 397–398, 1976.
- [8] T.-Y. Li and J. A. Yorke, "Period three implies chaos," *The American Mathematical Monthly*, vol. 82, no. 10, pp. 985–992, 1975.
- [9] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, "Geometry from a time series," *Physical Review Letters*, vol. 45, no. 9, pp. 712–716, 1980.
- [10] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick 1980 (Coventry, 1979/1980)*, D. A. Rand and L. S. Young, Eds., vol. 898 of *Lecture Notes in Mathematics*, pp. 366–381, Springer, Berlin, Germany, 1981.
- [11] R. Mañé, "On the dimension of the compact invariant sets of certain nonlinear maps," in *Dynamical Systems and Turbulence, Warwick 1980 (Coventry, 1979/1980)*, D. A. Rand and L. S. Young, Eds., vol. 898 of *Lecture Notes in Mathematics*, pp. 230–242, Springer, Berlin, Germany, 1981.
- [12] M. Casdagli, S. Eubank, J. D. Farmer, and J. Gibson, "State space reconstruction in the presence of noise," *Physica D*, vol. 51, no. 1–3, pp. 52–98, 1991.
- [13] T. Sauer, J. A. Yorke, and M. Casdagli, "Embedology," *Journal of Statistical Physics*, vol. 65, no. 3–4, pp. 579–616, 1991.
- [14] J. D. Farmer, E. Ott, and J. A. Yorke, "The dimension of chaotic attractors," *Physica D*, vol. 7, no. 1–3, pp. 153–180, 1983.
- [15] J. Theiler, "Estimating fractal dimension," *Journal of the Optical Society of America A*, vol. 7, no. 6, pp. 1055–1073, 1990.
- [16] V. I. Oseledec, "A multiplicative ergodic theorem. Characteristic Ljapunov exponents of dynamical systems," *Transactions of the Moscow Mathematical Society*, vol. 19, pp. 179–210, 1968.
- [17] A. N. Kolmogorov, "A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces," *Doklady Akademii Nauk SSSR*, vol. 119, no. 5, pp. 861–864, 1958.
- [18] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. Sh. Tsimring, "The analysis of observed chaotic data in physical systems," *Reviews of Modern Physics*, vol. 65, no. 4, pp. 1331–1392, 1993.
- [19] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, UK, 2nd edition, 2004.

- [20] M. Casdagli, D. Des Jardins, S. Eubank, et al., "Nonlinear modeling of chaotic time series: theory and applications," in *Applied Chaos*, J. H. Kim and J. Stringer, Eds., Wiley-Interscience Publication, chapter 15, pp. 335–380, John Wiley & Sons, New York, NY, USA, 1992.
- [21] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, Prentice-Hall, Englewood Cliffs, NJ, USA, 3rd edition, 1994.
- [22] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, NJ, USA, 2nd edition, 1999.
- [23] R. J. G. B. Campello, G. Favier, and W. C. do Amaral, "Optimal expansions of discrete-time Volterra models using Laguerre functions," *Automatica*, vol. 40, no. 5, pp. 815–822, 2004.
- [24] S. A. Billings, "Identification of nonlinear systems—a survey," *IEE Proceedings D*, vol. 127, no. 6, pp. 272–285, 1980.
- [25] I. J. Leontaritis and S. A. Billings, "Input-output parametric models for nonlinear systems—part II: stochastic nonlinear systems," *International Journal of Control*, vol. 41, no. 2, pp. 329–344, 1985.
- [26] S. A. Billings and Q. M. Zhu, "Rational model identification using an extended least-squares algorithm," *International Journal of Control*, vol. 54, no. 3, pp. 529–546, 1991.
- [27] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Transactions on Neural Networks*, vol. 1, no. 1, pp. 4–27, 1990.
- [28] S. Chen, C. F. N. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Transactions on Neural Networks*, vol. 2, no. 2, pp. 302–309, 1991.
- [29] L. Cao, Y. Hong, H. Fang, and G. He, "Predicting chaotic time series with wavelet networks," *Physica D*, vol. 85, no. 1-2, pp. 225–238, 1995.
- [30] L. Z. Guo, S. A. Billings, and D. Q. Zhu, "An extended orthogonal forward regression algorithm for system identification using entropy," *International Journal of Control*, vol. 81, no. 4, pp. 690–699, 2008.
- [31] S. A. Billings and H. L. Wei, "An adaptive orthogonal search algorithm for model subset selection and non-linear system identification," *International Journal of Control*, vol. 81, no. 5, pp. 714–724, 2008.
- [32] H.-L. Wei and S. A. Billings, "Model structure selection using an integrated forward orthogonal search algorithm assisted by squared correlation and mutual information," *International Journal of Modelling, Identification and Control*, vol. 3, no. 4, pp. 341–356, 2008.
- [33] L. Piroddi, "Simulation error minimisation methods for NARX model identification," *International Journal of Modelling, Identification and Control*, vol. 3, no. 4, pp. 392–403, 2008.
- [34] X. Hong, R. J. Mitchell, S. Chen, C. J. Harris, K. Li, and G. W. Irwin, "Model selection approaches for non-linear system identification: a review," *International Journal of Systems Science*, vol. 39, no. 10, pp. 925–946, 2008.
- [35] F. P. Pach, A. Gyenesei, and J. Abonyi, "MOSSFARM: model structure selection by fuzzy association rule mining," *Journal of Intelligent and Fuzzy Systems*, vol. 19, no. 6, pp. 399–407, 2008.
- [36] J. D. Farmer and J. J. Sidorowich, "Predicting chaotic time series," *Physical Review Letters*, vol. 59, no. 8, pp. 845–848, 1987.
- [37] J. P. Crutchfield and B. S. McNamara, "Equations of motion from a data series," *Complex Systems*, vol. 1, no. 3, pp. 417–452, 1987.
- [38] D. S. Broomhead and D. Lowe, "Multivariable functional interpolation and adaptive networks," *Complex Systems*, vol. 2, no. 3, pp. 321–355, 1988.
- [39] M. Casdagli, "Nonlinear prediction of chaotic time series," *Physica D*, vol. 35, no. 3, pp. 335–356, 1989.
- [40] J. Cremers and A. Hübler, "Construction of differential equations from experimental data," *Zeitschrift für Naturforschung A*, vol. 42, no. 8, pp. 797–802, 1987.
- [41] J. D. Farmer and J. J. Sidorowich, "Exploiting chaos to predict the future and reduce noise," in *Evolution, Learning and Cognition*, Y. C. Lee, Ed., pp. 277–330, World Scientific, Teaneck, NJ, USA, 1988.
- [42] M. Casdagli, "Chaos and deterministic versus stochastic nonlinear modelling," *Journal of the Royal Statistical Society. Series B*, vol. 54, no. 2, pp. 303–328, 1992.
- [43] H. D. I. Abarbanel, R. Brown, and J. B. Kadtke, "Prediction in chaotic nonlinear systems: methods for time series with broadband Fourier spectra," *Physical Review A*, vol. 41, no. 4, pp. 1782–1807, 1990.
- [44] M. B. Kennel and S. Isabelle, "Method to distinguish possible chaos from colored noise and to determine embedding parameters," *Physical Review A*, vol. 46, no. 6, pp. 3111–3118, 1992.
- [45] P. S. Linsay, "An efficient method of forecasting chaotic time series using linear interpolation," *Physics Letters A*, vol. 153, no. 6-7, pp. 353–356, 1991.
- [46] G. Xiaofeng and C. H. Lai, "Improvement of the local prediction of chaotic time series," *Physical Review E*, vol. 60, no. 5, pp. 5463–5468, 1999.

- [47] L. Cao, A. Mees, and K. Judd, "Dynamics from multivariate time series," *Physica D*, vol. 121, no. 1-2, pp. 75–88, 1998.
- [48] P. Grassberger, T. Schreiber, and C. Schaffrath, "Nonlinear time sequence analysis," *International Journal of Bifurcation and Chaos*, vol. 1, no. 3, pp. 521–547, 1991.
- [49] D. Kugiumtzis, O. C. Lingjærde, and N. Christophersen, "Regularized local linear prediction of chaotic time series," *Physica D*, vol. 112, no. 3-4, pp. 344–360, 1998.
- [50] L. Cao and A. S. Soofi, "Nonlinear deterministic forecasting of daily dollar exchange rates," *International Journal of Forecasting*, vol. 15, no. 4, pp. 421–430, 1999.
- [51] A. Lapedes and R. Farber, "Nonlinear signal processing using neural networks: prediction and system modelling," submitted to Proceedings of the IEEE, Los Alamos Report LA-UR 87-2662, 1987.
- [52] U. Parlitz, A. Hornstein, D. Engster, et al., "Identification of pre-sliding friction dynamics," *Chaos*, vol. 14, no. 2, pp. 420–430, 2004.
- [53] J.-S. Zhang and X.-C. Xiao, "Predicting chaotic time series using recurrent neural network," *Chinese Physics Letters*, vol. 17, no. 2, pp. 88–90, 2000.
- [54] J.-S. Zhang and X.-C. Xiao, "Predicting hyper-chaotic time series using adaptive higher-order nonlinear filter," *Chinese Physics Letters*, vol. 18, no. 3, pp. 337–340, 2001.
- [55] J. L. Breden and A. Hübler, "Reconstructing equations of motion from experimental data with unobserved variables," *Physical Review A*, vol. 42, no. 10, pp. 5817–5826, 1990.
- [56] J. L. Breden, F. Dinkelacker, and A. Hübler, "Noise in the modeling and control of dynamical systems," *Physical Review A*, vol. 42, no. 10, pp. 5827–5836, 1990.
- [57] G. Gouesbet, "Reconstruction of the vector fields of continuous dynamical systems from numerical scalar time series," *Physical Review A*, vol. 43, no. 10, pp. 5321–5331, 1991.
- [58] G. Gouesbet, "Reconstruction of vector fields: the case of the Lorenz system," *Physical Review A*, vol. 46, no. 4, pp. 1784–1796, 1992.
- [59] G. Gouesbet and J. Maquet, "Construction of phenomenological models from numerical scalar time series," *Physica D*, vol. 58, no. 1–4, pp. 202–215, 1992.
- [60] G. Gouesbet and C. Letellier, "Global vector-field reconstruction by using a multivariate polynomial  $L_2$  approximation on nets," *Physical Review E*, vol. 49, no. 6, pp. 4955–4972, 1994.
- [61] M. Giona, F. Lentini, and V. Cimagalli, "Functional reconstruction and local prediction of chaotic time series," *Physical Review A*, vol. 44, no. 6, pp. 3496–3502, 1991.
- [62] R. Brown, N. F. Rulkov, and E. R. Tracy, "Modeling and synchronizing chaotic systems from time-series data," *Physical Review E*, vol. 49, no. 5, pp. 3784–3800, 1994.
- [63] L. Le Sceller, C. Letellier, and G. Gouesbet, "Structure selection for global vector field reconstruction by using the identification of fixed points," *Physical Review E*, vol. 60, no. 2, pp. 1600–1606, 1999.
- [64] L. A. Aguirre, U. S. Freitas, C. Letellier, and J. Maquet, "Structure-selection techniques applied to continuous-time nonlinear models," *Physica D*, vol. 158, no. 1–4, pp. 1–18, 2001.
- [65] B. P. Bezruchko and D. A. Smirnov, "Constructing nonautonomous differential equations from experimental time series," *Physical Review E*, vol. 63, no. 1, Article ID 016207, 7 pages, 2000.
- [66] C. S. M. Lainscek, C. Letellier, and F. Schürer, "Ansatz library for global modeling with a structure selection," *Physical Review E*, vol. 64, no. 1, Article ID 016206, 15 pages, 2001.
- [67] M.-A. Boiron and J.-M. Malasoma, "Modélisation globale d'un système électrochimique," in *Comptes-Rendus de la 8ème Rencontre du Non-Linéaire, Non-Linéaire*, Orsay, France, 2005.
- [68] M.-A. Boiron and J.-M. Malasoma, "Modélisation globale d'un système électrochimique," in *Proceedings of the 5th Colloque on Temporal Chaos and Spatiotemporal Chaos*, Le Havre, France, December 2005.
- [69] C. Letellier, L. Le Sceller, E. Maréchal, et al., "Global vector field reconstruction from a chaotic experimental signal in copper electrodisolution," *Physical Review E*, vol. 51, no. 5, pp. 4262–4266, 1995.
- [70] A. D. Irving and T. Dewson, "Determining mixed linear-nonlinear coupled differential equations from multivariate discrete time series sequences," *Physica D*, vol. 102, no. 1-2, pp. 15–36, 1997.
- [71] C. Letellier, G. Gouesbet, and N. F. Rulkov, "Topological analysis of chaos in equivariant electronic circuits," *International Journal of Bifurcation and Chaos*, vol. 6, no. 12B, pp. 2531–2555, 1996.
- [72] C. Letellier, L. Le Sceller, G. Gouesbet, F. Lusseyran, A. Kemoun, and B. Izrar, "Recovering deterministic behavior from experimental time series in mixing reactor," *AIChE Journal*, vol. 43, no. 9, pp. 2194–2202, 1997.
- [73] J. Timmer, H. Rust, W. Horbelt, and H. U. Voss, "Parametric, nonparametric and parametric modelling of a chaotic circuit time series," *Physics Letters A*, vol. 274, no. 3-4, pp. 123–134, 2000.

- [74] W. Horbelt, J. Timmer, M. J. Bünner, R. Meucci, and M. Ciofini, "Identifying physical properties of a CO<sub>2</sub> laser by dynamical modeling of measured time series," *Physical Review E*, vol. 64, no. 1, Article ID 016222, 7 pages, 2001.
- [75] K. H. Chon and R. J. Cohen, "Linear and nonlinear ARMA model parameter estimation using an artificial neural network," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 3, pp. 168–174, 1997.
- [76] S. Haykin and J. Príncipe, "Making sense of a complex world," *IEEE Signal Processing Magazine*, vol. 15, no. 3, pp. 66–81, 1998.
- [77] L. A. Aguirre and E. C. Furtado, "Building dynamical models from data and prior knowledge: the case of the first period-doubling bifurcation," *Physical Review E*, vol. 76, no. 4, Article ID 046219, 13 pages, 2007.
- [78] A. M. Albano, A. Passamante, T. Hediger, and M. E. Farrell, "Using neural nets to look for chaos," *Physica D*, vol. 58, no. 1–4, pp. 1–9, 1992.
- [79] J. B. Elsner, "Predicting time series using a neural network as a method of distinguishing chaos from noise," *Journal of Physics A*, vol. 25, no. 4, pp. 843–850, 1992.
- [80] K. Aihara, T. Takabe, and M. Toyoda, "Chaotic neural networks," *Physics Letters A*, vol. 144, no. 6–7, pp. 333–340, 1990.
- [81] R. Bakker, J. C. Schouten, C. L. Giles, F. Takens, and C. M. van den Bleek, "Learning chaotic attractors by neural networks," *Neural Computation*, vol. 12, no. 10, pp. 2355–2383, 2000.
- [82] H. M. Henrique, E. L. Lima, and J. C. Pinto, "A bifurcation study on neural network models for nonlinear dynamic systems," *Latin American Applied Research*, vol. 28, no. 3, pp. 187–200, 1998.
- [83] R. A. Adomaitis, R. M. Farber, J. L. Hudson, I. G. Kevrekidis, M. Kube, and A. S. Lapedes, "Application of neural nets to system identification and bifurcation analysis of real world experimental data," in *Neural Networks: Biological Computers or Electronic Brains*, pp. 87–97, Springer, Paris, France, 1990.
- [84] R. Gençay and T. Liu, "Nonlinear modeling and prediction with feedforward and recurrent networks," *Physica D*, vol. 108, no. 1–2, pp. 119–134, 1997.
- [85] G. Boudjema and B. Cazelles, "Extraction of nonlinear dynamics from short and noisy time series," *Chaos, Solitons & Fractals*, vol. 12, no. 11, pp. 2051–2069, 2001.
- [86] H. J. Kim and K. S. Chang, "A method of model validation for chaotic chemical reaction systems based on neural networks," *Korean Journal of Chemical Engineering*, vol. 18, no. 5, pp. 623–629, 2001.
- [87] M. Small, *Applied Nonlinear Time Series Analysis*, vol. 52 of *World Scientific Series on Nonlinear Science. Series A*, World Scientific, Hackensack, NJ, USA, 2005.
- [88] L. A. Smith, "Identification and prediction of low-dimensional dynamics," *Physica D*, vol. 58, no. 1–4, pp. 50–76, 1992.
- [89] S. Ogawa, T. Ikeguchi, T. Matozaki, and K. Aihara, "Nonlinear modeling by radial basis function networks," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E79-A, no. 10, pp. 1608–1617, 1996.
- [90] T. Miyano, S. Kimoto, H. Shibuta, K. Nakashima, Y. Ikenaga, and K. Aihara, "Time series analysis and prediction on complex dynamical behavior observed in a blast furnace," *Physica D*, vol. 135, no. 3–4, pp. 305–330, 2000.
- [91] B. Pilgram, K. Judd, and A. Mees, "Modelling the dynamics of nonlinear time series using canonical variate analysis," *Physica D*, vol. 170, no. 2, pp. 103–117, 2002.
- [92] E. Bagarinao Jr., T. Nomura, K. Pakdaman, and S. Sato, "Generalized one-parameter bifurcation diagram reconstruction using time series," *Physica D*, vol. 124, no. 1–3, pp. 258–270, 1998.
- [93] E. Bagarinao Jr., K. Pakdaman, T. Nomura, and S. Sato, "Time series-based bifurcation diagram reconstruction," *Physica D*, vol. 130, no. 3–4, pp. 211–231, 1999.
- [94] E. Bagarinao Jr., K. Pakdaman, T. Nomura, and S. Sato, "Reconstructing bifurcation diagrams from noisy time series using nonlinear autoregressive models," *Physical Review E*, vol. 60, no. 1, pp. 1073–1076, 1999.
- [95] E. Bagarinao Jr., K. Pakdaman, T. Nomura, and S. Sato, "Reconstructing bifurcation diagrams of dynamical systems using measured time series," *Methods of Information in Medicine*, vol. 39, no. 2, pp. 146–149, 2000.
- [96] M. Barahona and C.-S. Poon, "Detection of nonlinear dynamics in short, noisy time series," *Nature*, vol. 381, no. 6579, pp. 215–217, 1996.
- [97] K. H. Chon, J. K. Kanters, R. J. Cohen, and N.-H. Holstein-Rathlou, "Detection of chaotic determinism in time series from randomly forced maps," *Physica D*, vol. 99, no. 4, pp. 471–486, 1997.

- [98] L. A. Aguirre, B. O. S. Teixeira, and L. A. B. Tôrres, "Using data-driven discrete-time models and the unscented Kalman filter to estimate unobserved variables of nonlinear systems," *Physical Review E*, vol. 72, no. 2, Article ID 026226, 12 pages, 2005.
- [99] N. Wessel, H. Malberg, R. Bauernschmitt, A. Schirdewan, and J. Kurths, "Nonlinear additive autoregressive model-based analysis of short-term heart rate variability," *Medical and Biological Engineering and Computing*, vol. 44, no. 4, pp. 321–330, 2006.
- [100] D. A. Vaccari and H.-K. Wang, "Multivariate polynomial regression for identification of chaotic time series," *Mathematical and Computer Modelling of Dynamical Systems*, vol. 13, no. 4, pp. 395–412, 2007.
- [101] D. J. de Oliveira, C. Letellier, M. E. D. Gomes, and L. A. Aguirre, "The use of synthetic input sequences in time series modeling," *Physics Letters A*, vol. 372, no. 32, pp. 5276–5282, 2008.
- [102] M. C. S. Coelho, E. M. A. M. Mendes, and L. A. Aguirre, "Testing for intracycle determinism in pseudoperiodic time series," *Chaos*, vol. 18, no. 2, Article ID 023125, p. 12, 2008.
- [103] M. Lei and G. Meng, "The influence of noise on nonlinear time series detection based on Volterra-Wiener-Korenberg model," *Chaos, Solitons & Fractals*, vol. 36, no. 2, pp. 512–516, 2008.
- [104] E. Floriani, T. Dudok de Wit, and P. Le Gal, "Nonlinear interactions in a rotating disk flow: from a Volterra model to the Ginzburg-Landau equation," *Chaos*, vol. 10, no. 4, pp. 834–847, 2000.
- [105] M. E. D. Gomes, A. V. P. Souza, H. N. Guimarães, and L. A. Aguirre, "Evidences of determinism in HRV signals," *Chaos*, vol. 10, no. 2, pp. 398–410, 2000.
- [106] L. A. Aguirre, C. Letellier, and J. Maquet, "Forecasting the time series of sunspot numbers," *Solar Physics*, vol. 249, no. 1, pp. 103–120, 2008.
- [107] S. A. Billings and S. Chen, "Identification of nonlinear rational systems using a prediction-error estimation algorithm," *International Journal of Systems Science*, vol. 20, no. 3, pp. 467–494, 1989.
- [108] Q. M. Zhu and S. A. Billings, "Recursive parameter estimation for nonlinear rational models," *Journal of Systems Engineering*, vol. 1, pp. 63–76, 1991.
- [109] Q. M. Zhu and S. A. Billings, "Parameter estimation for stochastic nonlinear rational models," *International Journal of Control*, vol. 57, no. 2, pp. 309–333, 1993.
- [110] M. V. Corrêa, L. A. Aguirre, and E. M. A. M. Menues, "Modeling chaotic dynamics with discrete nonlinear rational models," *International Journal of Bifurcation and Chaos*, vol. 10, no. 5, pp. 1019–1032, 2000.
- [111] O. Ménard, C. Letellier, and G. Gouesbet, "Map modeling by using rational functions," *Physical Review E*, vol. 62, no. 5, pp. 6325–6331, 2000.
- [112] D. Allingham, M. West, and A. I. Mees, "Wavelet reconstruction of nonlinear dynamics," *International Journal of Bifurcation and Chaos*, vol. 8, no. 11, pp. 2191–2201, 1998.
- [113] S. A. Billings and D. Coca, "Discrete wavelet models for identification and qualitative analysis of chaotic systems," *International Journal of Bifurcation and Chaos*, vol. 9, no. 7, pp. 1263–1284, 1999.
- [114] H. L. Wei and S. A. Billings, "Identification and reconstruction of chaotic systems using multiresolution wavelet decompositions," *International Journal of Systems Science*, vol. 35, no. 9, pp. 511–526, 2004.
- [115] R. J. G. B. Campello and W. C. do Amaral, "A relational approach for complex system identification," *Controle & Automação*, vol. 10, no. 3, pp. 139–148, 1999.
- [116] M. C. M. Teixeira and S. H. Zak, "Stabilizing controller design for uncertain nonlinear systems using fuzzy models," *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 2, pp. 133–142, 1999.
- [117] S. Guillaume, "Designing fuzzy inference systems from data: an interpretability-oriented review," *IEEE Transactions on Fuzzy Systems*, vol. 9, no. 3, pp. 426–443, 2001.
- [118] R. Ballini and F. Gomide, "Heuristic learning in recurrent neural fuzzy networks," *Journal of Intelligent & Fuzzy Systems*, vol. 13, no. 2–4, pp. 63–74, 2002.
- [119] H. L. Hiew and C. P. Tsang, "An adaptive fuzzy system for modeling chaos," *Information Sciences*, vol. 81, no. 3–4, pp. 193–212, 1994.
- [120] A. A. P. Santos, N. C. A. da Costa Jr., and L. dos Santos Coelho, "Computational intelligence approaches and linear models in case studies of forecasting exchange rates," *Expert Systems with Applications*, vol. 33, no. 4, pp. 816–823, 2007.
- [121] L. Le Sceller, C. Letellier, and G. Gouesbet, "Global vector field reconstruction including a control parameter dependence," *Physics Letters A*, vol. 211, no. 4, pp. 211–216, 1996.
- [122] M. Casdagli, "A dynamical systems approach to modeling input-output systems," in *Nonlinear Modeling and Forecasting*, M. Casdagli and S. Eubank, Eds., pp. 265–281, Addison Wesley, New York, NY, USA, 1992.

- [123] N. F. Hunter, "Application of nonlinear time-series models to driven systems," in *Nonlinear Modeling and Forecasting*, M. Casdagli and S. Eubank, Eds., pp. 467–491, Addison Wesley, New York, NY, USA, 1992.
- [124] I. J. Leontaritis and S. A. Billings, "Input-output parametric models for nonlinear systems—part I: deterministic non-linear systems," *International Journal of Control*, vol. 41, no. 2, pp. 303–328, 1985.
- [125] L. A. Aguirre and S. A. Billings, "Retrieving dynamical invariants from chaotic data using NARMAX models," *International Journal of Bifurcation and Chaos*, vol. 5, no. 2, pp. 449–474, 1995.
- [126] J. Lu, J. Lü, J. Xie, and G. Chen, "Reconstruction of the Lorenz and Chen systems with noisy observations," *Computers & Mathematics with Applications*, vol. 46, no. 8-9, pp. 1427–1434, 2003.
- [127] E. M. A. M. Mendes and S. A. Billings, "A note on discretization of nonlinear differential equations," *Chaos*, vol. 12, no. 1, pp. 66–71, 2002.
- [128] E. M. A. M. Mendes and C. Letellier, "Displacement in the parameter space versus spurious solution of discretization with large time step," *Journal of Physics A*, vol. 37, no. 4, pp. 1203–1218, 2004.
- [129] R. Reed, "Pruning algorithms—a survey," *IEEE Transactions on Neural Networks*, vol. 4, no. 5, pp. 740–747, 1993.
- [130] A. C. Tsoia and A. Back, "Discrete time recurrent neural network architectures: a unifying review," *Neurocomputing*, vol. 15, no. 3-4, pp. 183–223, 1997.
- [131] R. de Albuquerque Teixeira, A. P. Braga, R. H. C. Takahashi, and R. R. Saldanha, "Improving generalization of MLPs with multi-objective optimization," *Neurocomputing*, vol. 35, no. 1–4, pp. 189–194, 2000.
- [132] M. Small and C. K. Tse, "Minimum description length neural networks for time series prediction," *Physical Review E*, vol. 66, no. 6, Article ID 066701, 12 pages, 2002.
- [133] D. A. G. Vieira, J. A. Vasconcelos, and W. M. Caminhas, "Controlling the parallel layer perceptron complexity using a multiobjective learning algorithm," *Neural Computing & Applications*, vol. 16, no. 4-5, pp. 317–325, 2007.
- [134] B. Feil, J. Abonyi, and F. Szeifert, "Model order selection of nonlinear input-output models—a clustering based approach," *Journal of Process Control*, vol. 14, no. 6, pp. 593–602, 2004.
- [135] J. Madár, J. Abonyi, and F. Szeifert, "Genetic programming for the identification of nonlinear input-output models," *Industrial and Engineering Chemistry Research*, vol. 44, no. 9, pp. 3178–3186, 2005.
- [136] E. Baake, M. Baake, H. G. Bock, and K. M. Briggs, "Fitting ordinary differential equations to chaotic data," *Physical Review A*, vol. 45, no. 8, pp. 5524–5529, 1992.
- [137] G. L. Baker, J. P. Gollub, and J. A. Blackburn, "Inverting chaos: extracting system parameters from experimental data," *Chaos*, vol. 6, no. 4, pp. 528–533, 1996.
- [138] V. S. Anishchenko, A. N. Pavlov, and N. B. Janson, "Global reconstruction in the presence of *a priori* information," *Chaos, Solitons & Fractals*, vol. 9, no. 8, pp. 1267–1278, 1998.
- [139] G. Rowlands and J. C. Sprott, "Extraction of dynamical equations from chaotic data," *Physica D*, vol. 58, no. 1–4, pp. 251–259, 1992.
- [140] A. I. Mees, "Parsimonious dynamical reconstruction," *International Journal of Bifurcation and Chaos*, vol. 3, no. 3, pp. 669–675, 1993.
- [141] J. B. Kadtko, J. Brush, and J. Holzfuss, "Global dynamical equations and Lyapunov exponents from noisy chaotic time series," *International Journal of Bifurcation and Chaos*, vol. 3, no. 3, pp. 607–616, 1993.
- [142] L. A. Aguirre and S. A. Billings, "Dynamical effects of overparametrization in nonlinear models," *Physica D*, vol. 80, no. 1-2, pp. 26–40, 1995.
- [143] E. M. A. M. Mendes, *Identification of nonlinear discrete systems with intelligent structure detection*, Ph.D. thesis, University of Sheffield, Sheffield, UK, 1995.
- [144] L. A. Aguirre and E. M. A. M. Mendes, "Global nonlinear polynomial models: structure, term clusters and fixed points," *International Journal of Bifurcation and Chaos*, vol. 6, no. 2, pp. 279–294, 1996.
- [145] E. M. A. M. Mendes and S. A. Billings, "On identifying global nonlinear discrete models from chaotic data," *International Journal of Bifurcation and Chaos*, vol. 7, no. 11, pp. 2593–2601, 1997.
- [146] E. M. A. M. Mendes and S. A. Billings, "On overparametrization of nonlinear discrete systems," *International Journal of Bifurcation and Chaos*, vol. 8, no. 3, pp. 535–556, 1998.
- [147] E. M. A. M. Mendes and S. A. Billings, "An alternative solution to the model structure selection problem," *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, vol. 31, no. 6, pp. 597–608, 2001.
- [148] L. A. Aguirre, "Some remarks on structure selection for nonlinear models," *International Journal of Bifurcation and Chaos*, vol. 4, no. 6, pp. 1707–1714, 1994.

- [149] L. A. Aguirre and S. A. Billings, "Improved structure selection for nonlinear models based on term clustering," *International Journal of Control*, vol. 62, no. 3, pp. 569–587, 1995.
- [150] R. Brown, V. In, and E. R. Tracy, "Parameter uncertainties in models of equivariant dynamical systems," *Physica D*, vol. 102, no. 3-4, pp. 208–226, 1997.
- [151] L. A. Aguirre, R. A. M. Lopes, G. F. V. Amaral, and C. Letellier, "Constraining the topology of neural networks to ensure dynamics with symmetry properties," *Physical Review E*, vol. 69, no. 2, Article ID 026701, 11 pages, 2004.
- [152] J. D. Bomberger and D. E. Seborg, "Determination of model order for NARX models directly from input-output data," *Journal of Process Control*, vol. 8, no. 5-6, pp. 459–468, 1998.
- [153] J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, vol. 15 of *Series in Computer Science*, World Scientific, Singapore, 1989.
- [154] K. Judd and A. Mees, "On selecting models for nonlinear time series," *Physica D*, vol. 82, no. 4, pp. 426–444, 1995.
- [155] S. A. Billings, S. Chen, and M. J. Korenberg, "Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator," *International Journal of Control*, vol. 49, no. 6, pp. 2157–2189, 1989.
- [156] L. A. Aguirre and S. A. Billings, "Identification of models for chaotic systems from noisy data: implications for performance and nonlinear filtering," *Physica D*, vol. 85, no. 1-2, pp. 239–258, 1995.
- [157] L. A. Aguirre, P. F. Donoso-Garcia, and R. Santos-Filho, "Use of a priori information in the identification of global nonlinear models—a case study using a buck converter," *IEEE Transactions on Circuits and Systems I*, vol. 47, no. 7, pp. 1081–1085, 2000.
- [158] C. Letellier and L. A. Aguirre, "Investigating nonlinear dynamics from time series: the influence of symmetries and the choice of observables," *Chaos*, vol. 12, no. 3, pp. 549–558, 2002.
- [159] C. Letellier, L. A. Aguirre, and J. Maquet, "How the choice of the observable may influence the analysis of nonlinear dynamical systems," *Communications in Nonlinear Science and Numerical Simulation*, vol. 11, no. 5, pp. 555–576, 2006.
- [160] C. J. Cellucci, A. M. Albano, and P. E. Rapp, "Comparative study of embedding methods," *Physical Review E*, vol. 67, no. 6, Article ID 066210, 13 pages, 2003.
- [161] D. A. Smirnov, B. P. Bezruchko, and Y. P. Seleznev, "Choice of dynamical variables for global reconstruction of model equations from time series," *Physical Review E*, vol. 65, no. 2, Article ID 026205, 7 pages, 2002.
- [162] K. Judd and A. Mees, "Embedding as a modeling problem," *Physica D*, vol. 120, no. 3-4, pp. 273–286, 1998.
- [163] U. Parlitz, "Estimating model parameters from time series by autosynchronization," *Physical Review Letters*, vol. 76, no. 8, pp. 1232–1235, 1996.
- [164] A. Maybhate and R. E. Amritkar, "Use of synchronization and adaptive control in parameter estimation from a time series," *Physical Review E*, vol. 59, no. 1, pp. 284–293, 1999.
- [165] C. Tao, Y. Zhang, G. Du, and J. J. Jiang, "Fitting model equations to time series using chaos synchronization," *Physics Letters A*, vol. 332, no. 3-4, pp. 197–206, 2004.
- [166] U. S. Freitas, E. E. N. Macau, and C. Grebogi, "Using geometric control and chaotic synchronization to estimate an unknown model parameter," *Physical Review E*, vol. 71, no. 4, Article ID 047203, 4 pages, 2005.
- [167] E. J. Kostelich, "Problems in estimating dynamics from data," *Physica D*, vol. 58, no. 1-4, pp. 138–152, 1992.
- [168] L. Jaeger and H. Kantz, "Unbiased reconstruction of the dynamics underlying a noisy chaotic time series," *Chaos*, vol. 6, no. 3, pp. 440–450, 1996.
- [169] H. U. Voss, J. Timmer, and J. Kurths, "Nonlinear dynamical system identification from uncertain and indirect measurements," *International Journal of Bifurcation and Chaos*, vol. 14, no. 6, pp. 1905–1933, 2004.
- [170] J. M. Fullana, M. Rossi, and S. Zaleski, "Parameter identification in noisy extended systems: a hydrodynamic case," *Physica D*, vol. 103, no. 1-4, pp. 564–575, 1997.
- [171] P. Perona, A. Porporato, and L. Ridolfi, "On the trajectory method for the reconstruction of differential equations from time series," *Nonlinear Dynamics*, vol. 23, no. 1, pp. 13–33, 2000.
- [172] P. Connally, K. Li, and G. W. Irwin, "Prediction- and simulation-error based perceptron training: solution space analysis and a novel combined training scheme," *Neurocomputing*, vol. 70, no. 4-6, pp. 819–827, 2007.



- [173] J. L. Breeden and N. P. Packard, "A learning algorithm for optimal representation of experimental data," *International Journal of Bifurcation and Chaos*, vol. 4, no. 2, pp. 311–326, 1994.
- [174] J. Timmer, "Modeling noisy time series: physiological tremor," *International Journal of Bifurcation and Chaos*, vol. 8, no. 7, pp. 1505–1516, 1998.
- [175] H. U. Voss, A. Schwache, J. Kurths, and F. Mitschke, "Equations of motion from chaotic data: a driven optical fiber ring resonator," *Physics Letters A*, vol. 256, no. 1, pp. 47–54, 1999.
- [176] B. P. Bezruchko, A. S. Karavaev, V. I. Ponomarenko, and M. D. Prokhorov, "Reconstruction of time-delay systems from chaotic time series," *Physical Review E*, vol. 64, no. 5, Article ID 056216, 6 pages, 2001.
- [177] D. A. Smirnov, V. S. Vlaskin, and V. I. Ponomarenko, "Estimation of parameters in one-dimensional maps from noisy chaotic time series," *Physics Letters A*, vol. 336, no. 6, pp. 448–458, 2005.
- [178] E. G. Nepomuceno, R. H. C. Takahashi, G. F. V. Amaral, and L. A. Aguirre, "Nonlinear identification using prior knowledge of fixed points: a multiobjective approach," *International Journal of Bifurcation and Chaos*, vol. 13, no. 5, pp. 1229–1246, 2003.
- [179] L. A. Aguirre, E. C. Furtado, and L. A. B. Tôrres, "Evaluation of dynamical models: dissipative synchronization and other techniques," *Physical Review E*, vol. 74, no. 6, Article ID 066203, 16 pages, 2006.
- [180] L. A. Aguirre and S. A. Billings, "Validating identified nonlinear models with chaotic dynamics," *International Journal of Bifurcation and Chaos*, vol. 4, no. 1, pp. 109–125, 1994.
- [181] E. Bagarinao and S. Sato, "Algorithm for vector autoregressive model parameter estimation using an orthogonalization procedure," *Annals of Biomedical Engineering*, vol. 30, no. 2, pp. 260–271, 2002.
- [182] A. Garulli, C. Mocenni, A. Vicino, and A. Tesi, "Integrating identification and qualitative analysis for the dynamic model of a lagoon," *International Journal of Bifurcation and Chaos*, vol. 13, no. 2, pp. 357–374, 2003.
- [183] K. Rodríguez-Vázquez and P. J. Fleming, "Evolution of mathematical models of chaotic systems based on multiobjective genetic programming," *Knowledge and Information Systems*, vol. 8, no. 2, pp. 235–256, 2005.
- [184] L. M. Pecora, T. L. Carroll, and J. F. Heagy, "Statistics for mathematical properties of maps between time series embeddings," *Physical Review E*, vol. 52, no. 4, pp. 3420–3439, 1995.
- [185] R. Brown, N. F. Rul'kov, and E. R. Tracy, "Modeling and synchronizing chaotic systems from experimental data," *Physics Letters A*, vol. 194, no. 1-2, pp. 71–76, 1994.
- [186] G. Gouesbet, S. Meunier-Guttin-Cluzel, and O. Ménard, "Global reconstructions of equations of motion from data series, and validation techniques, a review," in *Chaos and Its Reconstruction*, G. Gouesbet, S. Meunier-Guttin-Cluzel, and O. Ménard, Eds., pp. 1–160, Nova Science, New York, NY, USA, 2003.
- [187] D. Coca and S. A. Billings, "Identification of coupled map lattice models of complex spatio-temporal patterns," *Physics Letters A*, vol. 287, no. 1-2, pp. 65–73, 2001.
- [188] S. Ishii and M.-A. Sato, "Reconstruction of chaotic dynamics by on-line EM algorithm," *Neural Networks*, vol. 14, no. 9, pp. 1239–1256, 2001.
- [189] L. A. Aguirre and Á. V. P. Souza, "Stability analysis of sleep apnea time series using identified models: a case study," *Computers in Biology and Medicine*, vol. 34, no. 4, pp. 241–257, 2004.
- [190] J. Maquet, C. Letellier, and L. A. Aguirre, "Scalar modeling and analysis of a 3D biochemical reaction model," *Journal of Theoretical Biology*, vol. 228, no. 3, pp. 421–430, 2004.
- [191] O. Ménard, C. Letellier, J. Maquet, L. Le Sceller, and G. Gouesbet, "Analysis of a nonsynchronized sinusoidally driven dynamical system," *International Journal of Bifurcation and Chaos*, vol. 10, no. 7, pp. 1759–1772, 2000.
- [192] C. Letellier, L. Le Sceller, P. Dutertre, G. Gouesbet, Z. Fei, and J. L. Hudson, "Topological characterization and global vector field reconstruction of an experimental electrochemical system," *Journal of Physical Chemistry*, vol. 99, no. 18, pp. 7016–7027, 1995.
- [193] N. B. Tufillaro, P. Wyckoff, R. Brown, T. Schreiber, and T. Molteno, "Topological time-series analysis of a string experiment and its synchronized model," *Physical Review E*, vol. 51, no. 1, pp. 164–174, 1995.
- [194] C. Letellier and G. Gouesbet, "Topological characterization of reconstructed attractors modding out symmetries," *Journal de Physique II*, vol. 6, no. 11, pp. 1615–1638, 1996.
- [195] M. Small and K. Judd, "Comparisons of new nonlinear modeling techniques with applications to infant respiration," *Physica D*, vol. 117, no. 1–4, pp. 283–298, 1998.

- [196] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, "Testing for nonlinearity in time series: the method of surrogate data," *Physica D*, vol. 58, no. 1–4, pp. 77–94, 1992.
- [197] J. Timmer, "Power of surrogate data testing with respect to nonstationarity," *Physical Review E*, vol. 58, no. 4, pp. 5153–5156, 1998.
- [198] J. P. Barnard, C. Aldrich, and M. Gerber, "Identification of dynamic process systems with surrogate data methods," *AIChE Journal*, vol. 47, no. 9, pp. 2064–2075, 2001.
- [199] C. Letellier, T. D. Tsankov, G. Byrne, and R. Gilmore, "Large-scale structural reorganization of strange attractors," *Physical Review E*, vol. 72, no. 2, Article ID 026212, 12 pages, 2005.
- [200] P. E. McSharry and L. A. Smith, "Consistent nonlinear dynamics: identifying model inadequacy," *Physica D*, vol. 192, no. 1–2, pp. 1–22, 2004.
- [201] R. Brown, N. F. Rul'kov, and E. R. Tracy, "Modeling and synchronizing chaotic systems from time-series data," *Physical Review E*, vol. 49, no. 5, pp. 3784–3800, 1994.
- [202] C. Letellier, O. Ménard, and L. A. Aguirre, "Validation of selected global models," in *Modeling and Forecasting Financial Data: Techniques of Nonlinear Dynamics*, A. S. Soofi and L. Cao, Eds., pp. 283–302, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
- [203] C. Sarasola, F. J. Torrealdea, A. d'Anjou, and M. Graña, "Cost of synchronizing different chaotic systems," *Mathematics and Computers in Simulation*, vol. 58, no. 4–6, pp. 309–327, 2002.
- [204] R. Gilmore and M. Lefranc, *The Topology of Chaos: Alice in Stretch and Squeezeland*, Wiley-Interscience, New York, NY, USA, 2002.
- [205] C. Letellier, J. Maquet, L. Le Sceller, G. Gouesbet, and L. A. Aguirre, "On the non-equivalence of observables in phase-space reconstructions from recorded time series," *Journal of Physics A*, vol. 31, no. 39, pp. 7913–7927, 1998.
- [206] H. J. A. F. Tulleken, "Grey-box modelling and identification using physical knowledge and Bayesian techniques," *Automatica*, vol. 29, no. 2, pp. 285–308, 1993.
- [207] E. Eskinat, S. H. Johnson, and W. L. Luyben, "Use of auxiliary information in system identification," *Industrial and Engineering Chemistry Research*, vol. 32, no. 9, pp. 1981–1992, 1993.
- [208] T. A. Johansen, "Identification of non-linear systems using empirical data and prior knowledge—an optimization approach," *Automatica*, vol. 32, no. 3, pp. 337–356, 1996.
- [209] M. V. Corrêa, L. A. Aguirre, and R. R. Saldanha, "Using steady-state prior knowledge to constrain parameter estimates in nonlinear system identification," *IEEE Transactions on Circuits and Systems I*, vol. 49, no. 9, pp. 1376–1381, 2002.
- [210] L. A. Aguirre, M. F. S. Barroso, R. R. Saldanha, and E. M. A. M. Mendes, "Imposing steady-state performance on identified nonlinear polynomial models by means of constrained parameter estimation," *IEE Proceedings: Control Theory and Applications*, vol. 151, no. 2, pp. 174–179, 2004.
- [211] P. E. Rapp, T. I. Schmah, and A. I. Mees, "Models of knowing and the investigation of dynamical systems," *Physica D*, vol. 132, no. 1–2, pp. 133–149, 1999.
- [212] G. Byrne, R. Gilmore, and C. Letellier, "Distinguishing between folding and tearing mechanisms in strange attractors," *Physical Review E*, vol. 70, no. 5, Article ID 056214, 9 pages, 2004.
- [213] G. F. V. Amaral, C. Letellier, and L. A. Aguirre, "Piecewise affine models of chaotic attractors: the Rössler and Lorenz systems," *Chaos*, vol. 16, no. 1, Article ID 013115, 14 pages, 2006.
- [214] R. Gilmore, "Summary of the second workshop on measures of complexity and chaos," *International Journal of Bifurcation and Chaos*, vol. 3, no. 3, pp. 491–524, 1993.
- [215] R. M. May, "Deterministic models with chaotic dynamics," *Nature*, vol. 256, no. 5514, pp. 165–166, 1975.
- [216] R. M. May, "Simple mathematical models with very complicated dynamics," *Nature*, vol. 261, no. 5560, pp. 459–467, 1976.
- [217] M. Hénon, "A two-dimensional mapping with a strange attractor," *Communications in Mathematical Physics*, vol. 50, no. 1, pp. 69–77, 1976.
- [218] M. Xu, G. Chen, and Y.-T. Tian, "Identifying chaotic systems using Wiener and Hammerstein cascade models," *Mathematical and Computer Modelling*, vol. 33, no. 4–5, pp. 483–493, 2001.
- [219] K. Ikeda, "Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system," *Optics Communications*, vol. 30, no. 2, pp. 257–261, 1979.
- [220] O. E. Rössler, "An equation for continuous chaos," *Physics Letters A*, vol. 57, no. 5, pp. 397–398, 1976.
- [221] M. C. Mackey and L. Glass, "Oscillation and chaos in physiological control systems," *Science*, vol. 197, no. 4300, pp. 287–289, 1977.
- [222] B. van der Pol and M. van der Mark, "The heartbeat considered as a relaxation oscillation, and an electrical model of the heart," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, Series 7*, vol. 6, pp. 763–775, 1928, Pl 10–12.

- [223] N. A. Gershenfeld and A. A. Weigend, "The future of time series: learning and understanding," in *Time Series Prediction: Forecasting the Future and Understanding the Past*, A. A. Weigend and N. A. Gershenfeld, Eds., pp. 1–70, Addison-Wesley, New York, NY, USA, 1994.
- [224] D. R. Rigney, A. L. Goldberger, W. C. Ocasio, Y. Ichimaru, G. B. Moody, and R. G. Mark, "Multi-channel physiological data: description and analysis," in *Time Series Prediction: Forecasting the Future and Understanding the Past*, A. A. Weigend and N. A. Gershenfeld, Eds., pp. 105–129, Addison-Wesley, New York, NY, USA, 1994.
- [225] A. A. Weigend and N. A. Gershenfeld, *Time Series Prediction: Forecasting the Future and Understanding the Past*, Addison-Wesley, New York, NY, USA, 1994.
- [226] M. Casdagli and A. A. Weigend, "Exploring the continuum between deterministic and stochastic modeling," in *Time Series Prediction: Forecasting the Future and Understanding the Past*, A. A. Weigend and N. A. Gershenfeld, Eds., pp. 347–366, Addison-Wesley, New York, NY, USA, 1994.
- [227] D. Kaplan, "A geometrical statistic for detecting deterministic dynamics," in *Time Series Prediction: Forecasting the Future and Understanding the Past*, A. A. Weigend and N. A. Gershenfeld, Eds., pp. 415–428, Addison-Wesley, New York, NY, USA, 1994.
- [228] M. Paluš, "Detecting nonlinearity in multivariate time series," *Physics Letters A*, vol. 213, no. 3–4, pp. 138–147, 1996.
- [229] L. A. Aguirre, V. C. Barros, and Á. V. P. Souza, "Nonlinear multivariable modeling and analysis of sleep apnea time series," *Computers in Biology and Medicine*, vol. 29, no. 3, pp. 207–228, 1999.
- [230] L. Cao and A. Mees, "Deterministic structure in multichannel physiological data," *International Journal of Bifurcation and Chaos*, vol. 10, no. 12, pp. 2767–2780, 2000.
- [231] C. S. M. Lainscsek, F. Schürer, and J. B. Kadtko, "A general form for global dynamical data models for three-dimensional systems," *International Journal of Bifurcation and Chaos*, vol. 8, no. 5, pp. 899–914, 1998.
- [232] C. Letellier, L. A. Aguirre, J. Maquet, and R. Gilmore, "Evidence for low dimensional chaos in sunspot cycles," *Astronomy & Astrophysics*, vol. 449, no. 1, pp. 379–387, 2006.
- [233] J. Maquet, C. Letellier, and L. A. Aguirre, "Global models from the Canadian lynx cycles as a direct evidence for chaos in real ecosystems," *Journal of Mathematical Biology*, vol. 55, no. 1, pp. 21–39, 2007.
- [234] N. F. Rul'kov, A. R. Volkovskii, A. Rodríguez-Lozano, E. del Río, and M. G. Velarde, "Mutual synchronization of chaotic self-oscillators with dissipative coupling," *International Journal of Bifurcation and Chaos*, vol. 2, no. 3, pp. 669–676, 1992.
- [235] T. Matsumoto, L. O. Chua, and M. Komuro, "The double scroll," *IEEE Transactions on Circuits and Systems*, vol. 32, no. 8, pp. 797–818, 1985.
- [236] L. A. Aguirre, G. G. Rodrigues, and E. M. A. M. Mendes, "Nonlinear identification and cluster analysis of chaotic attractors from a real implementation of Chua's circuit," *International Journal of Bifurcation and Chaos*, vol. 7, no. 6, pp. 1411–1423, 1997.
- [237] B. Cannas, S. Cincotti, M. Marchesi, and F. Pilo, "Learning of Chua's circuit attractors by locally recurrent neural networks," *Chaos, Solitons & Fractals*, vol. 12, no. 11, pp. 2109–2115, 2001.