PERGAMON

Chaos, Solitons and Fractals 13 (2002) 1099–1113

CHAOS
SOLITONS & FRACTALS

www.elsevier.com/locate/chaos

Should all the species of a food chain be counted to investigate the global dynamics?

Christophe Letellier [a,*], Luis A. Aguirre [b], Jean Maquet [a], M.A. Aziz-Alaoui [c]

[a] *CORIA UMR 6614 – Université et INSA de Rouen, Place Emile Blondel, F-76821 Mont Saint-Aignan Cedex, France*
[b] *MACSIN, Dept. Enga. Eletrônica, Universidade Federal de Minas Gerais, Av. Antonio Carlos 6627, Belo Horizonte 31270-901, Brazil*
[c] *Laboratoire de Mathématiques, Fac. Sc. Tech., BP 540, F-76058 Le Havre Cedex, France*

Received 29 March 2001

Abstract

A fairly realistic three-species food-chain model based on Lotka–Volterra and Leslie–Gower schemes is investigated assuming that just a single scalar time series is available. The paper uses tools borrowed from the theory of nonlinear dynamical systems. The quality of the different phase portraits reconstructed is tested. Such a situation would arise in practice whenever only a single species is counted. It is found that the dynamical analysis can be safely performed when a single species involved in the food chain is counted if many thousands of observations are available. If not, a global model can be obtained from the available data and subsequently used to produce all the data required for a detailed analysis. In this case, however, the choice of which species to consider in order to obtain a model is crucially important.

1. Introduction

When an ecological model involving few species is investigated, complete knowledge of the states of the system may require the measurement of all the interacting species. Unfortunately, when a real ecological system is investigated, it is very often not possible or it becomes too expensive to count all the species involved. However, one of the most interesting concepts from nonlinear dynamical system theory is that the time evolution of a single species may be used to reconstruct a phase portrait equivalent to the original one. For instance, if the ecological model under study involves N species designated by a state vector $\boldsymbol{x}(t)$, the original phase portrait is embedded in a space spanned by all the components of $\boldsymbol{x}(t)$. The time evolution $\{\boldsymbol{x}(t)\}$ is the solution of the dynamical system $\dot{\boldsymbol{x}}(t) = \boldsymbol{\Phi}(\boldsymbol{x}(t))$, where $\boldsymbol{\Phi}$ designates the vector field. If we suppose that we can only measure one component, i.e. one scalar function $X = G(\boldsymbol{x}(t))$, it may be possible to reconstruct a phase portrait which is expected to be equivalent to the original one by using delay or derivative coordinates [1]. In such a situation, is the knowledge of the time evolution of one species enough to predict or to reproduce the time evolution of the whole food chain with a global model? This is the natural question that this paper will address for a fairly realistic three-species food-chain model which is based on Lotka and Leslie–Gower schemes as studied in [2–4]. It describes a prey population $\tilde{x}$ which serves as the only food for a predator $\tilde{y}$. This specialist predator $\tilde{y}$ is, in turn, the prey of a generalist predator $\tilde{z}$ which is assumed to reproduce sexually. In the last decade, a number of studies have been carried out to analyze ecosystem models based on Lotka–Volterra or Holling type-I schemes, see [5–9]. There is now a considerable literature on discrete and continuous models of this kind. Particularly, such papers often address the problem of existence and stability of equilibria or persistence of food chains. However, other interesting properties are the natures of the dynamical regimes exhibited by a dynamical system. Such an approach is

* Corresponding author.
E-mail address: christophe.letellier@coria.fr (C. Letellier).

0960-0779/02/$ - see front matter
PII: S0960-0779(01)00116-3

particularly important when the asymptotic behavior of the system is chaotic. In particular, recent studies have shown that chaotic dynamics may play an important role in continuous-time models for ecological systems and there is some evidence that the real time evolution of species involved in two or three food chains could be characterized by chaotic attractors as observed in many natural food chains, see [3–8] for example.

To determine whether the dynamics underlying a real ecological system is chaotic, it is rather important to possess measures of the time evolutions of the interacting species. Unfortunately, practical measurements pose concrete difficulties. First, it is impossible or too expensive to measure the whole state vector, i.e. to count simultaneously all the interacting species. Second, most algorithms (e.g. Lyapunov exponent and correlation dimension) assume lots of clean and stationary data. One way of avoiding such problems is to follow a two-step procedure: first obtain a global model from a short possibly noisy and weakly non-stationary time series. Then use the model to generate all the data required for analysis. This has been illustrated for a very large number of simulated cases in [10]. Nevertheless, it is not always possible to obtain a global model from a single time variable and such a possibility depends crucially on the time series available [11]. It is therefore important to determine which species should be counted to optimize the chance of obtaining a global model. This paper aims to address this last question.

The paper is organized as follows. Section 2 briefly introduces the realistic ecological model studied. Section 3 investigates the phase portrait induced by each species involved in the food chain, i.e. whether the food-chain dynamics can be safely investigated from a single time series. Section 4 is devoted to quantifying the observability of the system and to obtaining global models from the most observable variables. The main conclusions of the paper are summarized in Section 5.

2. Mathematical formulation of the model and equilibria

A quite realistic model involving a three-species food chain is considered. A typical situation of such a scheme would involve rodents, snakes and peacocks [12]. Interaction between the specialist predator $\tilde{y}$ and its prey $\tilde{x}$ may be modeled by a Volterra scheme, i.e. predator population dies out exponentially in the absence of its prey. Nevertheless, for a more realistic model, the interaction between this predator $\tilde{y}$ and the generalist predator $\tilde{z}$ is rather modeled by the Leslie–Gower scheme, where the loss in a predator population is proportional to the reciprocal of per capita availability of its most favorite food, as will be detailed in what follows.

2.1. The mathematical model

A mathematical model governing the system can be written as a set of three continuous autonomous first-order differential equations, [2] namely

$$\begin{aligned}
\dot{\tilde{x}} &= a_1\tilde{x} - b_1\tilde{x}^2 - \frac{\omega_0\tilde{x}\tilde{y}}{\tilde{x}+d_0},\\
\dot{\tilde{y}} &= -a_2\tilde{y} + \frac{\omega_1\tilde{x}\tilde{y}}{\tilde{x}+d_1} - \frac{\omega_2\tilde{y}\tilde{z}}{\tilde{y}+d_2},\\
\dot{\tilde{z}} &= c_0\tilde{z}^2 - \frac{\omega_3\tilde{z}^2}{\tilde{y}+d_3},
\end{aligned} \tag{1}$$

where the dot indicates derivatives with respect to time (or time derivatives), $\tilde{x}$ is the density of prey at the bottom of the food chain, $\tilde{y}$ is the density of the specialist predator, and $\tilde{z}$ is the density of the generalist predator.

The meanings of the parameters are given below:

- a_1 is the rate of the self-growth for prey $\tilde{x}$;
- a_2 measures the rate at which $\tilde{y}$ will die out when there is no prey left;
- ω_i's are the maximum values attainable by each per capita rate;
- d_0 and d_1 quantify the extent to which the environment provides protection to the prey $\tilde{x}$;
- b_1 measures the strength of competition among prey individuals $\tilde{x}$;
- d_2 is the value of $\tilde{y}$ at which per capita removal rate of $\tilde{y}$ becomes $\omega_2/2$;
- d_3 represents the residual loss in $\tilde{z}$ population due to severe scarcity of its favorite food $\tilde{y}$;
- c_0 describes the growth rate of the generalist predator $\tilde{z}$ by sexual reproduction, the number of males and females being assumed to be equal.

These parameters assume only positive values. The total number of parameters of this model may be reduced from 12 to 8 by using a scaling transformation that reads as [4]

$$x = \frac{a_1}{b_1}\tilde{x}, \qquad y = \frac{a_1^2}{b_1\omega_0}\tilde{y}, \qquad z = \frac{a_1^3}{b_1\omega_0\omega_2}\tilde{z}, \qquad t = \frac{\tilde{t}}{a_1}. \tag{2}$$

The rescaled system then becomes

$$\begin{aligned} \dot{x} &= x(1-x) - \frac{xy}{x+a}, \\ \dot{y} &= -by + \frac{cxy}{x+d} - \frac{yz}{y+e}, \\ \dot{z} &= fz^2 - \frac{gz^2}{y+h}, \end{aligned} \tag{3}$$

where the bifurcation parameters are defined as

$$\begin{aligned} a &= \frac{b_1 d_0}{a_1}, \qquad b = \frac{a_2}{a_1}, \qquad c = \frac{\omega_1}{a_1}, \qquad d = \frac{b_1 d_1}{a_1}, \\ e &= \frac{d_2\omega_0 b_1}{a_1^2}, \qquad f = \frac{c_0 a_1^2}{b_1\omega_0\omega_2}, \qquad g = \frac{\omega_3}{\omega_2}, \qquad h = \frac{d_3\omega_0 b_1}{a_1^2}. \end{aligned} \tag{4}$$

In [3] and [4] it was shown that this system displays an impressive breadth of complex phenomena, including chaos. It has been found that this continuous-time ecosystem model can produce a type-I intermittency, homoclinic orbits or multistability. The existence and the stability of the fixed points have been studied. Setting $\dot{x} = \dot{y} = \dot{z} = 0$ in Eq. 3, the equilibria are obtained on solving the resulting equations [4]. Six fixed points are identified. One of them, F_0, is located at the origin of the phase space $\mathbb{R}^3(x, y, z)$. The others are

$$F_1(1,0,0), \qquad F_2\left(0, \frac{g}{f} - h, b\left(h - e - \frac{g}{f}\right)\right), \qquad F_3\left(\frac{bd}{c-b}, (1-x_3)(a+x_3), 0\right)$$

and

$$F_{4\pm} = \begin{vmatrix} x_{4\pm} = \frac{1-a}{2} \pm \sqrt{\frac{(a+1)^2}{4} - \frac{g}{f} + h}, \\ y_{4\pm} = \frac{g}{f} - h, \\ z_{4\pm} = \left(h - e - \frac{g}{f}\right)\left(b - \frac{x_{4\pm}c}{x_{4\pm}+d}\right). \end{vmatrix} \tag{5}$$

Fixed points F_0 and F_1 always exist. On the other hand, it can be shown that F_2 never exists in the positive octant where the realistic model is defined. The existence of the fixed points $F_{4\pm}$ have been studied by investigating their local and global stability [4].

In this study, the bifurcation parameters are chosen to correspond to a realistic scheme [2,13]. They read as

$$\begin{aligned} &a_2 = 1.0, \quad b_1 = 0.06, \quad d_0 = 10.0, \quad d_1 = 10.0, \quad d_2 = 10.0, \quad d_3 = 20.0, \\ &\omega_0 = 1.0, \quad \omega_1 = 2.0, \quad \omega_2 = 0.405, \quad \omega_3 = 1.0, \quad c_0 = 0.03. \end{aligned} \tag{6}$$

The parameter values are chosen on the basis of previous studies [2] and correspond to quantitative measures of attributes of the rodent–snake–peacock food chain. Fig. 1 shows plane projections of the chaotic attractor obtained for $a_1 = 1.93$.

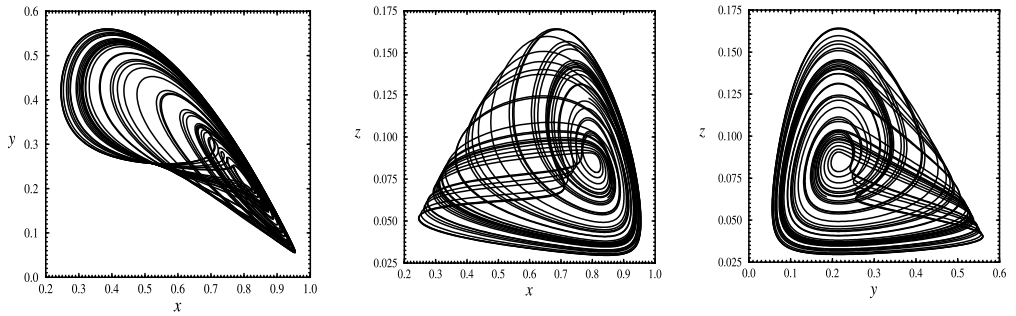


Fig. 1. Projections of the chaotic attractor onto the three different planes of the phase space $\mathbb{R}^3(x, y, z)$ for $a_1 = 1.93$.

3. Analysis of the food chain through a single observed (or counted) species: embedology

3.1. Generalities

When one would like to investigate a multi-species food chain, one faces the important question: *Do I have to count all the species within a period of time* (*that is to say: do I have to actually know the time evolution of all the species that compose my studied food chain*)? This situation with wide ramifications is connected with the important practical problem of analyzing a nonlinear dynamical system from a single time series. When real dynamics are investigated from a single scalar time series, the first step of any dynamical analysis is to reconstruct a phase portrait which is expected to be equivalent to the original, but not measured, phase portrait. In our case, the "measured" time series will be constituted by the time evolution of only one of the three species involved in the food-chain system (3), i.e. either x, y or z. The reconstructed phase portrait is thus embedded in a space spanned by delay coordinates

$$\{X(t), X(t+\tau), X(t+2\tau), \ldots, X(t+(d_E-1)\tau)\}$$

or the derivative coordinates

$$\left\{X(t), \dot{X}(t), \ddot{X}(t), \ldots, \frac{\mathrm{d}^{d_E-1}X(t)}{\mathrm{d}t^{d_E-1}}\right\},$$

which have been shown to be equivalent [14].

Takens [15] has proposed and analyzed these different techniques for constructing a vector of whatever dimension is desired from a time series of a single scalar in the ideal case, i.e. no noise and infinite data. The most interesting result from nonlinear dynamical system theory is that the time evolution of the system in the phase space does not necessarily require knowledge of all the dynamical variables involved in the complete description of a state of the system because, according to the information redundancy principle, a time series corresponding to the time evolution of a single species may be sufficient to provide the relevant information to investigate a many-species food chain.

Starting from a single scalar time series, it is possible to reconstruct a phase portrait which is expected to be equivalent to the original phase portrait. The best equivalence which may be expected is when the map which transforms the original phase portrait into the reconstructed one is an embedding defined as follows.

Definition 1. A one-to-one continuous function from a compact set to $\mathbb{R}^d$ is called an *embedding* of the set (or a topological embedding). The number d is called the *embedding dimension* of the set.

A remarkable theorem by Takens [15] states:

Theorem 1. *A finite-dimensional attractor can always be embedded in some* $\mathbb{R}^d$; *the sufficient embedding dimension* d *can be only a unit greater than twice the dimension of the attractor*.

When using a single time series, one may be sure of having obtained the best quality of reconstruction when a diffeormorphism between the reconstructed state portrait and the original one is found. Such a quality may be obtained when the embedding dimension d is sufficiently large. According to Takens' theorem, it is sufficient to take d greater than or equal to $2D_{\mathrm{H}}+1$ where D_{H} ideally refers to the Hausdorff dimension which could be conveniently approximated by the correlation dimension estimated by Grassberger and Procaccia's algorithm [16]. Nevertheless, a dimension less than d may be sufficient to obtain an equivalent phase portrait. Such a dimension may be estimated by using the false nearest neighbors [17], for which an improved algorithm has been recently proposed by Cao [18]. The latter is used here.

Theoretically, i.e. with an infinite amount of data without any noise, all the dynamical variables are equivalent and may be used. Unfortunately, when one faces data from the real world, necessarily corrupted by noise and discretized in time, such equivalence is not always observed and we may find that some variables are better than others [11]. Such a difficulty may become more severe when only poor-quality data are available, as often happens in the study of ecological systems [19], where an uncontrolled nonlinear superimposition of fluctuations may be present [20]. It is therefore rather important to state whether the three species may be equivalently used to investigate the dynamics of the whole food chain.

3.2. Numerical experiments

As previously explained, the interesting feature of derivative (or delay) coordinates is that they can span a reconstructed phase portrait which is expected to be equivalent to the original phase portrait. In order to do so, the time series

itself is plotted versus its time derivatives. The phase portraits embedded in the space spanned by the derivative coordinates (X, Y, Z) are called the *differential embedding*. For the three variables of the three-species food-chain system (3), the dimension of the differential embedding is obviously equal to 3 (Fig. 2). The reconstructed state space is therefore spanned by the three variables reading as

$$\begin{aligned} X &= x, y \text{ or } z, \\ Y &= \dot{X}, \\ Z &= \ddot{X}. \end{aligned} \tag{7}$$

Plane projections of the three induced phase portraits are displayed in Fig. 3.

A first-return map is then computed for each reconstructed phase portrait. The Poincaré sections P_i are defined as follows:

$$\mathscr{P}_x = \{(X_n, Z_n) \in \mathbb{R}^2 |\ X_n > 0.80,\ Y_n = 0,\ Z_n < 0\} \tag{8}$$

when $X = x$,

$$\mathscr{P}_y = \{(X_n, Z_n) \in \mathbb{R}^2 |\ X_n < 0.21,\ Y_n = 0,\ Z_n > 0\} \tag{9}$$

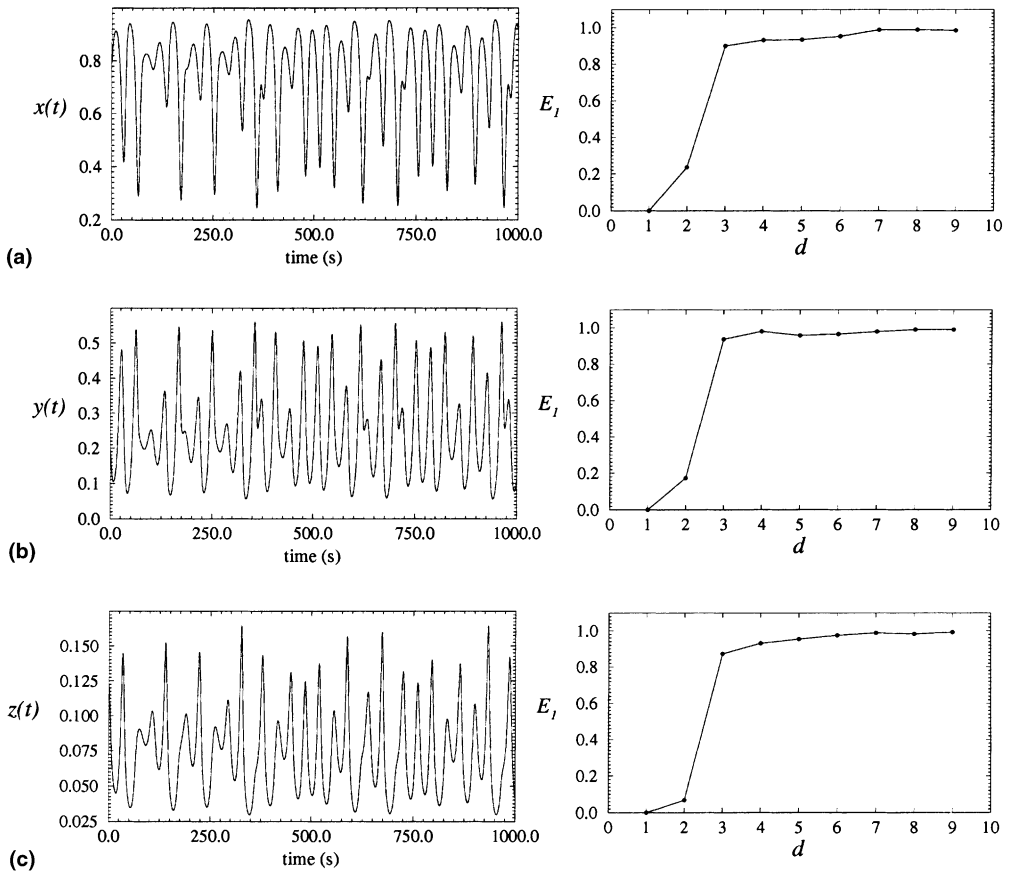


Fig. 2. The three dynamical variables (x, y or z) can be used for investigating the dynamical behavior of the three-species food-chain model. For each variable, the embedding dimension is equal to 3. In the method proposed by Cao, the ordinates saturates when the embedding dimension is large enough. Index $E_1(d)$ measures the relative change in the average distance between two neighboring points in $\mathbb{R}^d$ and their respective images in $\mathbb{R}^{d+1}$ when the embedding dimension is increased from d to $d+1$. The embedding dimension are computed by using a phase space reconstructed with the delay coordinates with a time delay τ equal to $20\delta t = 0.3$ s. (a) x-variable; (b) y-variable; (c) z-variable.

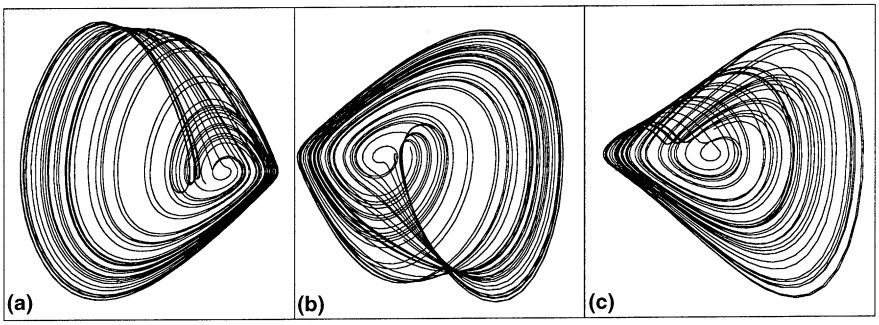


Fig. 3. XY-plane projections of the different differential embeddings induced by the three dynamical variables of the system (3): they give a fairly faithful visual representation of the original attractor shown in Fig. 1. (a) x-variable; (b) y-variable; (c) z-variable.

when $X = y$ and

$$\mathscr{P}_z = \{(X_n, Z_n) \in \mathbb{R}^2 |\ Y_n = 0,\ Z_n > 0\} \tag{10}$$

when $X = z$. The three first-return maps are found to be constituted by three monotonic branches (Figs. 4(a)–(c)) as observed for the original attractor (Fig. 4(d)). The population of periodic orbits is then extracted and is the same as the one from the original attractor reported in Table 1.

We showed that the evolution of the orbit spectrum is rather similar to those of the Rössler system, at least until no more than three monotonic branches are identified on the first-return map. It is then interesting to check whether the relative organizations of the periodic orbits in the phase space are similar too. Since the model considered here is tridimensional, its periodic orbits, viewed as knots, can be characterized by linking numbers [21]. Using the concept of

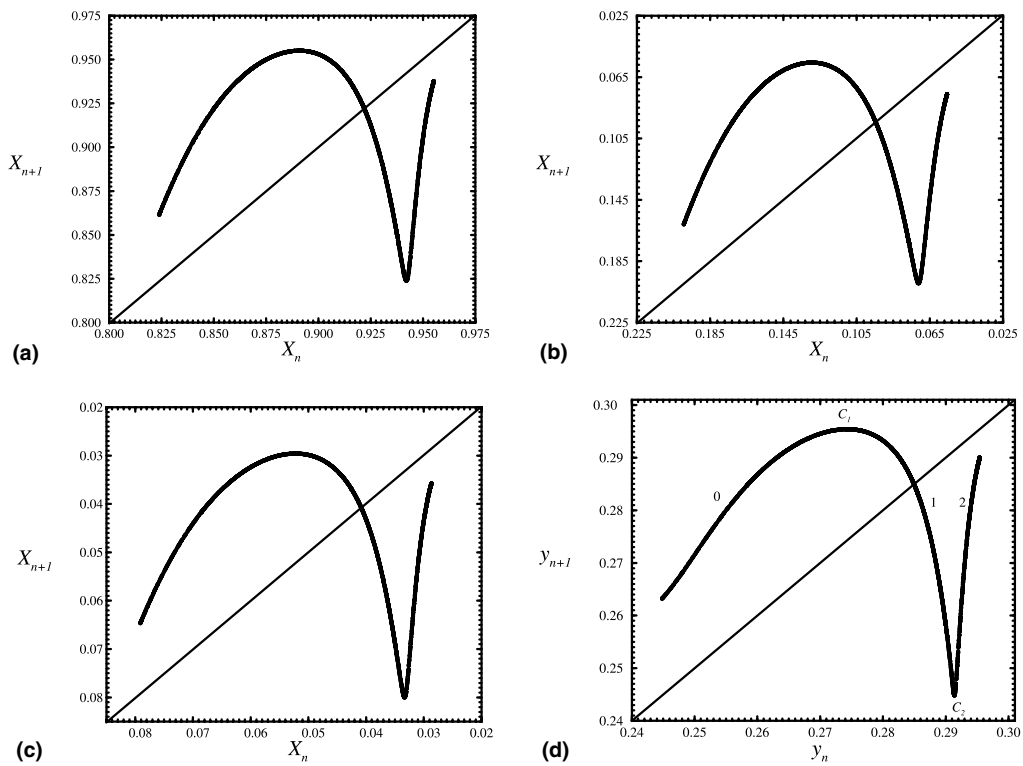


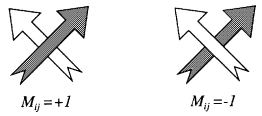
Fig. 4. The first-return map to the Poincaré section $\mathscr{P}$ is found to be constituted by three monotonic branches separated by two critical points C_1 and C_2. (a) x-induced portrait; (b) y-induced portrait; (c) z-induced portrait; (d) original phase portrait.

Table 1
Population of periodic orbits embedded within the attractor for the a_1 bifurcation parameter equal to 1.93 ($c_0 = 0.03$)

1	201	21 100	211
10	20 110	21 101	210
1011	20 111	211 011	210 211
10 110	201 110	211 020	21 021
101	20 100	211 021	21 020
100	20 200	2110	210 201
100 101	20	21 110	21 010
200 101	21	211 100	21 011
200 100	21 201	21 120	2111
200	21 200	21 121	211 121

knot-holder introduced by Birman and Williams [22], a template synthesizing the topological properties of the attractor may then be built [21]. A template is a branched manifold. To each monotonic branch of the first-return map corresponds a strip on the template. Thus, the template of the chaotic attractor generated by the three-species model (3) will be constituted by three strips for $a_1 = 1.93$. An important property of the strips is their local torsion, i.e. the number of half-turns of the tangent space counted when a trajectory makes one revolution on the chaotic attractor through a given strip. From the first-return map, the parity of the local torsion may be defined. An increasing branch is associated with an even number of half-turns; it is an order-preserving branch. On the contrary, a decreasing branch is an order-reversing branch and is associated with a strip whose local torsion is odd. In our case, the template will be constituted by two even and one odd strips according to the first-return map.

The chaotic attractor generated by the three-species food-chain model is characterized by the template displayed in Fig. 5. Such a template may be defined by a linking matrix M_{ij} whose diagonal elements M_{ii} describe the number of half-turns of the ith strip and off-diagonal elements M_{ij} provide the number of crossings between the ith and the jth branches [23] oriented according to the following convention:



In the case of the three-species system, the linking matrix of the three-strip attractor reads as

$$M_{ij} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & +1 & +1 \\ 0 & +1 & +2 \end{bmatrix}. \tag{11}$$

Such a template allows one to predict all the linking numbers between couples of periodic orbits. Indeed, all of them have been found in agreement with those counted on plane projections of the corresponding periodic orbits extracted

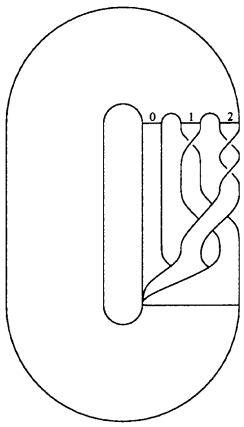


Fig. 5. Template of the three-strip chaotic attractor generated by the three-species model.

from the attractor. In order to do so, each oriented crossing $\chi(p)$ is identified on the plane projection and oriented with the aid of the third direction of the phase space according to the usual convention.

Linking numbers are also computed for each reconstructed attractor. All of them are well predicted by the template characterizing the original attractor (Fig. 5). Details of topological characterization by templates can be found in [21,24,25]. The three reconstructed phase portraits are thus topologically equivalent to the original one. The dynamics is therefore preserved under the reconstruction process. One may therefore investigate this three-species food chain by recording the time evolution of any single species. Since the topological properties have been checked for the three-strip attractor, the equivalence is also checked for all attractors whose populations of periodic orbits are a subset of the one extracted for this attractor. Indeed, linking numbers are unchanged while the corresponding periodic orbits are not involved in a bifurcation.

4. Observability indices

When a real food chain is investigated from counting a single species, one could attempt to obtain a global model from this time series using global modeling techniques. Thus a set of ordinary differential equations or difference equations obtained would be able to reproduce the time evolution of the counted species. Few algorithms have been proposed in the literature. In this paper, two different modeling techniques are used, the so-called nonlinear autoregressive moving average (NARMA) technique as introduced by Leontaritis and Billings [26] and the global modeling technique providing a set of ordinary differential equations as developed by Gouesbet and Maquet [27]. However, since global modeling is much more sensitive to the choice of the observed variable, i.e. the species counted, than topological analysis, discussed in the previous section, the observability of the dynamics of the food chain here investigated from the species is first quantified using an observability index originally defined in [28] and applied in the study of nonlinear dynamics in [11].

In order to quantify the observability, the indices introduced in [11] will be used. The concept of observability in linear systems theory is standard [29]. Consider the system

$$\begin{aligned} \dot{\boldsymbol{x}} &= A\boldsymbol{x} + B\boldsymbol{u}, \\ s &= C\boldsymbol{x}, \end{aligned} \tag{12}$$

where $\boldsymbol{x} \in \mathbb{R}^m$ is the state vector, $s \in \mathbb{R}^r$ is the measured vector, $\boldsymbol{u} \in \mathbb{R}^p$ is the input vector and $\{A, B, C\}$ are constant matrices. In this paper, $r = 1$ and s is the observable. In the case of nonlinear systems, A is replaced by the jacobian matrix evaluated along the trajectory. In other words, given an orbit, the system jacobian matrix is estimated at each point for which the observability indices (see below) are calculated. Subsequently, the average along the trajectory is computed. Matrix C, on the other hand, is trivially obtained from the measurement function designated by h. The system (12) is said to be state observable at time t_f if the initial state $\boldsymbol{x}(0)$ can be uniquely determined from knowledge of a finite time history of the input $\boldsymbol{u}(t)$ and output $s(t)$, $0 \leqslant t \leqslant t_f$ [29]. It should be noted that the definition is still valid for autonomous systems, that is, for $\boldsymbol{u}(t) = 0$ or $B = 0$.

Definition 2. The system (12) is said to be state observable if the *observability matrix*

$$Q = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-r} \end{bmatrix} \tag{13}$$

is of full rank, that is if $\operatorname{rank}(Q) = m$.

This definition is a "yes" or "no" measurement of observability, i.e. the system is either observable or not. In practice, however, a system may gradually become unobservable as a parameter is varied or, for nonlinear systems, it seems reasonable to suppose that there might be regions in state space that are less observable than others. Hence it is useful to define the observability index as [28]

$$\delta = \frac{|\lambda_{\min}(QQ^{\mathrm{T}})|}{|\lambda_{\max}(QQ^{\mathrm{T}})|}, \tag{14}$$

where $(\cdot)^{\mathrm{T}}$ indicates the transpose, and $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalues. Then $0 \leqslant \delta \leqslant 1$, and the lower bound is reached when the system is unobservable. It should be noticed that the index (14) is a type of condition number of the observability matrix.

To compute the observability indices, the starting point is to obtain matrix A seen in Eq. (12). Because such an equation is linear, the system should be linearized along a trajectory $(\tilde{x}, \tilde{y}, \tilde{z})$ in such a way that at each point we have a constant matrix

$$A = \left[\begin{array}{ccc} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \\ \dfrac{\partial f_3}{\partial x} & \dfrac{\partial f_3}{\partial y} & \dfrac{\partial f_3}{\partial z} \end{array} \right]_{x=\tilde{x},\ y=\tilde{y},\ z=\tilde{z}} \tag{15}$$

that describes the local dynamics.

The second step is to define the observable through the measurement function and choose C accordingly. For instance, if we measure the observable $s = x$, then $C = [1\ 0\ 0]$. If we measure $s = y + 0.5z$, then $C = [0\ 1\ 0.5]$, and so on. In these cases, because we use only one observable, s is a scalar and $r = 1$ in (13). The formation of matrix Q (see (13)) follows directly and the computation of the observability index using (14) can be performed using standard software. In this way, at each point on the trajectory we have a value for the observability indices. In this section we have always shown the values of such indices averaged along the trajectory considered.

The average observability indices were computed for the three dynamical variables of the ecological model (3). It was found that

$$\delta_x = 0.0736, \qquad \delta_y = 0.0019, \qquad \delta_z = 2.4299 \times 10^{-5}. \tag{16}$$

It seems that the dynamics underlying this three-species food chain is definitely less observable from the observations of the generalist predator than from the prey or the specialist predator. The very low observability of the dynamics from the z-variable is quite surprising because, from a visual inspection of the time series shown in Fig. 2, it is not obvious that the variables may not be equivalent for observing the dynamics. From these observability indices, it appears that counting the generalist predator, located at the end of the food chain, is not an adequate strategy for investigating the dynamics of this food chain. It is much more efficient to count the prey or the specialist predator.

5. Global modeling from the different species

Consider a continuous-time dynamical system described by a set of ordinary differential equations

$$\dot{\boldsymbol{x}} = \boldsymbol{f}\ (\boldsymbol{x}; \boldsymbol{\mu}), \tag{17}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the state vector that depends on a parameter t called the time and $\boldsymbol{f}$, the so-called vector field, is an n-component smooth function generating a flow ϕ_t. Also, $\boldsymbol{\mu} \in \mathbb{R}^p$ is the parameter vector with p components, assumed to be constant in this work.

It is assumed that a single variable is measured, and it is desired to obtain a dynamical model, from that single time series and with no prior knowledge, that will represent the original dynamics in some sense. In the remainder of this section, two different model representations will be described. Such representations will be used on simulated and real data to illustrate the main ideas of this work.

5.1. Discrete-time models

Here it is considered that the counted species is $y(k) = h(\boldsymbol{x}(kT_s))$, $k = 0, 1, \ldots$, where $\boldsymbol{x}$ is the state vector and T_s is the time interval between two observations (to count the number of individuals of a given species). In many cases, the time evolution of the observed $y(k)$ can be described by a NARMA model [26] of the form

$$y(k) = F^{\ell}[y(k-1), \ldots, y(k-n_y), e(k), \ldots, e(k-n_e)], \tag{18}$$

where n_y and n_e are the maximum lags considered for the process and noise terms, respectively. Moreover, $y(k)$ is the output time series and $e(k)$ accounts for uncertainties, possible noise and unmodeled dynamics. $F^{\ell}[\cdot]$ is some nonlinear function of $y(k)$ and $e(k)$.

In this paper, the map $F^{\ell}[\cdot]$ is a polynomial of degree $\ell \in \mathbb{Z}^{+}$. In order to estimate the parameters of this map, Eq. (18) can be expressed as

$$y(k) = \boldsymbol{\psi}(k-1)^{\mathrm{T}}\hat{\boldsymbol{\theta}} + \xi(k), \tag{19}$$

where $\xi(k)$ are the identification residuals. Moreover, $\boldsymbol{\psi}(k-1)$ is a vector which contains output and residual terms up to and including time $k-1$ and $\hat{\boldsymbol{\theta}}$ is the estimated parameter vector obtained by minimizing the following cost function [30]:

$$J_N(\hat{\boldsymbol{\theta}}) = \frac{1}{N}\sum_{k=1}^{N} \xi^2(k, \hat{\boldsymbol{\theta}}). \tag{20}$$

Parameter estimation is usually performed for a linear-in-the-parameters orthogonal model which is closely related to (19) and which is represented as

$$y(k) = \sum_{i=1}^{n_{\mathrm{p}}+n_{\xi}} g_i w_i(k) + \xi(k), \tag{21}$$

where $n_{\mathrm{p}} + n_{\xi}$ is the number of (process plus noise) terms in the model, $\{g_i\}_{i=1}^{n_{\mathrm{p}}+n_{\xi}}$ are parameters and the monomials $\{w_i(k)\}_{i=1}^{n_{\mathrm{p}}+n_{\xi}}$ are orthogonal over the data records. Finally, parameters of the model in Eq. (19) can be calculated from the $\{g_i\}_{i=1}^{n_{\mathrm{p}}+n_{\xi}}$. This procedure has two major advantages, namely: (i) it reduces inaccuracies due to numerical ill-conditioning; (ii) it aids in selecting the structure of the final model.

A criterion for selecting the most important terms in the model can be devised as a byproduct of the orthogonal parameter estimation procedure. The reduction in the MSPE due to the inclusion of the ith term, $g_i w_i(k)$, in the auxiliary model of Eq. (21) is $(1/N)g_i^2\overline{w_i^2(k)}$. Expressing this reduction in terms of the total MSPE yields the error reduction ratio (ERR):

$$[\mathrm{ERR}]_i \doteq \frac{g_i^2\overline{w_i^2(k)}}{\overline{y^2(k)}}, \quad i = 1, 2, \ldots, n_{\mathrm{p}} + n_{\xi}. \tag{22}$$

Hence those terms with large values of ERR are selected to form the model.

In order to search for discrete-time models the following procedure was used. Data obtained by integrating the model (3) were sampled according to the procedure detailed in [31]; this yielded $T_{\mathrm{s}} = 1.0$. Time series with 600 (this corresponds roughly to 15 cycles only) observations were used in every case. A set with 296 candidate terms was generated by taking all the linear, quadratic and cubic combinations up to delay $n_y = 10$. A moving average (MA) model with 10 linear terms was used during parameter estimation in order to reduce bias. Such terms are not used in the simulations and therefore are not shown below. The ERR criterion, previously described, was used to classify the candidate terms according to their relative importance in explaining the underlying dynamics. The choice of how many terms to retain was made based upon topological analysis of the identified candidate models. This entire procedure was repeated in three different situations, namely assuming only the x variable was available, and then, in turn, the same was performed using the y and z time series. Several models were found when using x and y time series, whereas no model was found from the z-variable. This is totally in accord with the observability indices.

Two models obtained from the x and y time series are, respectively,

$$\begin{aligned} x(k) = {} & 3.38040x(k-1) - 4.30812x(k-2) + 2.56162x(k-3) - 1.06161x(k-4) - 1.21955x^2(k-5) \\ & + 2.56978x(k-1)x(k-5)x(k-6) - 3.26196x(k-3)x(k-4)x(k-6) + 0.48632x(k-5) \\ & + 2.53047x^2(k-4)x(k-5) + 0.80920x(k-4)x(k-7) - 4.55223\times 10^{-3}x^2(k-1)x(k-8) \\ & + 1.47483x(k-3)x(k-6) - 0.23716x^2(k-5)x(k-6) - 0.74444x(k-1)x(k-7) - 0.45312x^2(k-6) \\ & + 0.50283x^2(k-2)x(k-3) - 2.02429x(k-1)x(k-4)x(k-5), \end{aligned} \tag{23}$$

$$\begin{aligned} y(k) = {} & 2.66466y(k-1) - 2.98381y(k-2) + 2.13113y(k-3) - 0.88570y(k-4) + 0.12437y(k-5) \\ & + 3.59558\times 10^{-3}y(k-6) + 1.68574y^2(k-1) + 0.46715y(k-2)y^2(k-3) + 1.89658y^2(k-1)y(k-5) \\ & - 1.00584\times 10^{-3} + 0.99980y^3(k-4) + 0.96611y(k-1)y^2(k-6) - 0.99972y^2(k-5)y(k-6) \\ & - 3.31664y(k-1)y(k-3) - 0.19571y(k-1)y(k-4)y(k-6) + 1.28748y(k-2)y(k-4) \\ & - 2.61306y(k-2)y(k-4)y(k-6). \end{aligned} \tag{24}$$

These model terms are presented in the same order of priority as determined by the ERR criterion. Only the auto-regressive parts of the models are shown.

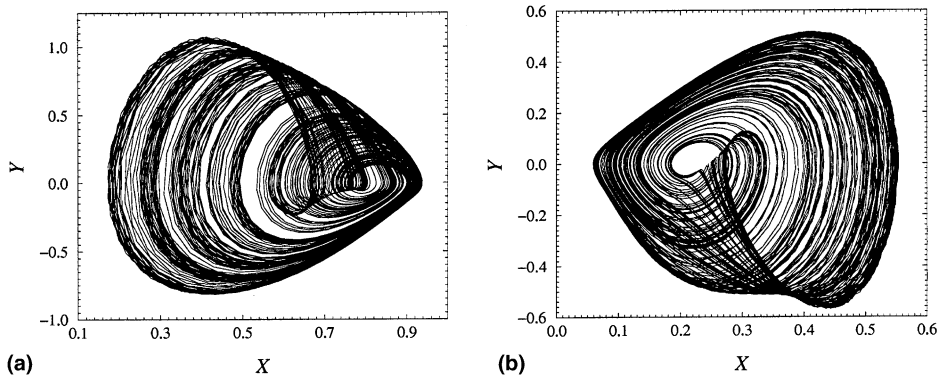


Fig. 6. Plane projection of the attractors generated by the models obtained from the (a) x- and (b) y-variables, respectively. No model from the z-variable – the generalist predator – has been obtained with the NARMA technique. These two plots should be compared to Figs. 3(a) and 3(b), respectively.

As expected from the observability indices, it is very difficult to obtain a global model from the time evolution of the generalist predator. In fact, no model has been obtained from such a time series. Successful models were only obtained from the prey and the specialist predator time series. The corresponding attractors are shown in Fig. 6. From a visual inspection, these two models seem to generate phase portraits rather equivalent to those expected. Compare the model attractor from the x-variable (Fig. 6(a)) with the phase portrait reconstructed from the same variable (Fig. 3(a)) and the model attractor from the y-variable (Fig. 6(b)) with the phase portrait shown in Fig. 3(b). A more accurate validation is performed by computing a first-return map to the associated Poincaré sections $\mathscr{P}_x$ and $\mathscr{P}_y$, respectively. The first-return map associated with the model attractor obtained from the x-variable is slightly different than the original (Fig. 4(a)). In particular, a layered structure is observed on the first increasing branch. It is also clear that the branches are less developed than the original ones, as is easily exhibited when the population of periodic orbits are extracted. A less numerous population of periodic orbits is extracted. It may depend slightly on the time series used for extracting periodic orbits (3500 cycles are used here). Some periodic orbits may be very poorly represented (or even not completely present) in this time series. Nevertheless, the population extracted from the attractor generated by the model obtained from the y-variable is closer to the population extracted from the original attractor (Table 1). Indeed, the model from the x-variable shares 12 orbits with the original system whereas the model from the y-variable includes 23 orbits (Table 1). Moreover, no layered structure is observed (Fig. 7(b)) as for the original first-return map (Fig. 4(b)).

The model obtained from the time evolution of the specialist is therefore much better than the one obtained from counting the prey. Consequently, it appears that if one has to choose one species to count, the specialist predator should be preferred. This may be understood from a dynamical point of view since it is the single species that directly interacts with the other two.

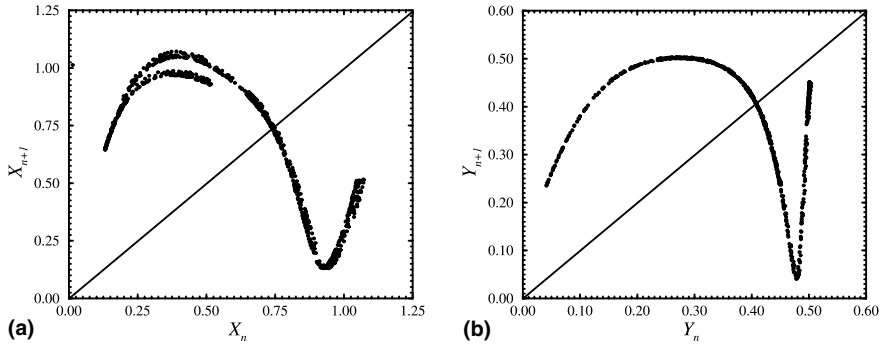


Fig. 7. First-return maps to a Poincaré section for the two models obtained from the (a) x- and (b) y-variables, respectively. To generate these plots, models (23) and (24) were iterated to produce 30×10^3 data points. On the other hand, such models were obtained from just 600 observations of the respective species. These plots should be compared to Figs. 4(a) and 4(b), respectively.

5.2. Continuous-time models

Very similar results are obtained with a global modeling technique using continuous-time models. In this approach it is assumed that a single scalar time series, $X_1 = x = h(\boldsymbol{x})$, is recorded, where, as before, $h(\cdot)$ is a measurement function. The aim is then to obtain a vector field equivalent to the original system using a basis consisting of the observable and its derivatives such as

$$\begin{aligned} \dot{X}_1 &= X_2, \\ \dot{X}_2 &= X_3, \\ &\vdots \\ \dot{X}_{d_e} &= F(X_1, X_2, \ldots, X_{d_e}), \end{aligned} \tag{25}$$

where d_e is the embedding dimension and the model function F depends on d_e variables which are s and the $d_e - 1$ successive derivatives of s. The model function F can be estimated by using a multivariate polynomial basis on nets [32]. The algorithm requires the definition of modeling parameters, which are:

1. d_e, the embedding dimension,
2. N_c, the number of centers at which the function is evaluated,
3. Δt, the time step between two successive centers; in this work Δt is constant, but this is not a requirement,
4. N_p, the number of retained multinomials and
5. τ_w, the window length on which the derivatives are computed by using polynomial interpolation over the window.

Derivatives are then obtained by analytically deriving such polynomials. The estimated model function, $\hat{F}$, then reads as

$$\hat{F}(X_1, X_2, \ldots, X_{d_e}) = \sum_{p=1}^{N_p} \theta_p \psi^p, \tag{26}$$

where θ_p are the parameters and ψ^p are multivariate monomials (or multinomials) of the form

$$\psi^p = X_1^{n_1} X_2^{n_2} \cdots X_{d_e}^{n_{d_e}}, \tag{27}$$

where the integers p are related to n_{d_e}-tuplets $(n_1, n_2, \ldots, n_{d_e})$ by a bijective relationship discussed in [32]. The modeling parameters d_e, N_c, Δt, N_p and τ_w can be determined with the aid of an error function described in [33]. The model (26) is based on a differential embedding, i.e. a phase space spanned by the derivative coordinates.

Writing (26) at N_c centers on the data yields a set of N_c equations of the form

$$\begin{bmatrix} \dot{X}_{d_e}(1) \\ \vdots \\ \dot{X}_{d_e}(N_c) \end{bmatrix} = \begin{bmatrix} \psi^{r1}(1) & \cdots & \psi^{rn_p}(1) \\ \vdots & & \vdots \\ \psi^{r1}(N_c) & \cdots & \psi^{rn_p}(N_c) \end{bmatrix} \begin{bmatrix} \theta_{r1} \\ \vdots \\ \theta_{rn_p} \end{bmatrix}, \tag{28}$$

where $ri \neq rj$ and $1 \leqslant r1, \ldots, rn_p \leqslant N_p$ and the numbers inside the parentheses indicate to which center the variable is related. Once the model structure is determined the parameter vector $[\hat{\theta}_{r1} \cdots \hat{\theta}_{rn_p}]^T$ can be estimated by standard least-squares techniques.

Three continuous differential models are obtained with this technique. Indeed, a model has been obtained from each variable with the modeling parameters as follows. The embedding dimension d_e is equal to 3 for all of them as shown in Fig. 2. The integration step is $\delta t = 0.01$ of the dimensionless time. For all the time series, the window τ_w for interpolating the time series is $7\delta t$. The three remaining modeling parameters $(N_c, \Delta t, N_p)$ are

- $(200, 17, 35)$ when $s = x$;
- $(110, 20, 56)$ when $s = y$;
- $(300, 20, 56)$ when $s = z$.

Thus the model from the x-variable has 35 terms and, consequently, has an order-4 polynomial model function. The other two models have 56 terms which constitute order-5 polynomial model functions. Plane projection of the attractor generated by these three models are shown in Fig. 8. The two models obtained from the x- and y-variables are stable and may be numerically integrated over more than 3500 cycles. Contrary to this, the best model obtained from the z-variable is unstable although its transient regime evolves very close to an apparently chaotic attractor. After that transient regime, the trajectory is ejected to infinity (Fig. 8(c)). Roughly 80 cycles can only be generated. This second modeling

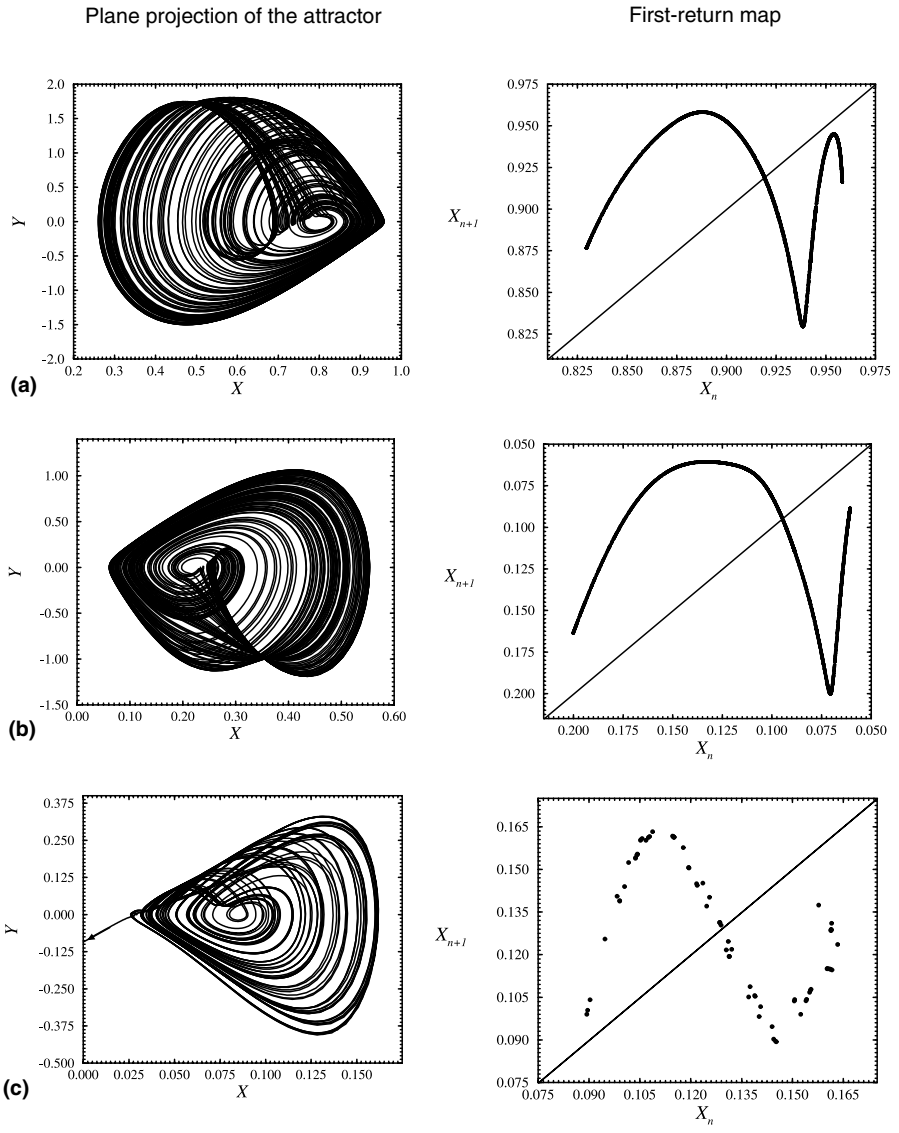


Fig. 8. Plane projection of the attractors generated by the models obtained from the three variables of the ecological model (a) x, (b) y and (c) z, respectively. The first-return maps are computed in the Poincaré sections $\mathscr{P}_x$, $\mathscr{P}_y$ and $\mathscr{P}_z$, respectively. Only an unstable model is obtained from the z-variable. After a metastable chaotic regime, the trajectory is ejected to infinity as indicated by the arrow. Only 80 cycles are obtained to compute the first-return map, which is thicker than the other two.

technique also confirms that the z-variable is not a good observable for the dynamics, i.e. a global model is very difficult to obtain from it.

The quality of the models obtained from the x- and y-variables is investigated using first-return maps to Poincaré sections, as for the discrete-time models. The model from the x-variable presents a four-monotonic-branch first-return map, i.e. the dynamics is slightly more developed than the original attractor. As a consequence, the population of orbits with a period less than 7 is larger for this model (63 orbits) than for the original attractor (40 orbits reported in Table 1). The model attractor from the x-variable is therefore characterized by a four-strip template. Among the four strips, three have the same relative organization in the phase space as the original attractor. Thus, the corresponding template is the one shown in Fig. 5 but with one additional strip having a local torsion equal to +3. This model therefore generates an attractor which corresponds to the dynamics that the original system could generate but with slightly different control parameters [3]. Comparing with the model obtained with the NARMA technique, the quality of the model from the

x-variable is roughly the same since the continuous-time model shares only 13 orbits with the original system. The NARMA technique tends to reduce the number of periodic orbits whereas the modeling technique with derivative coordinates tends to increase it.

As observed with the discrete-time models, the model obtained from the y-variable is definitely the best one. Indeed, the first-return map (Fig. 8(b)) is very similar to the one computed with the original system (Fig. 4(b)). In that case, the model shares 27 orbits with the original system (3). A similar conclusion can therefore be given, i.e. counting the specialist predators provides the best possible data.

6. Conclusion

A fairly realistic three-species food-chain model based on the couplings between a Lotka–Volterra and a Leslie–Gower scheme has been investigated. It has been shown that such a food chain can be investigated by only counting a single species. Indeed, the phase space reconstruction techniques allow one to obtain a representation of the phase space which is equivalent to the original one. In other words, dynamical analysis in a phase space reconstructed from the time evolution of a single species can be safely performed. The time evolution of the non-counted species are contained in the recorded one because all the species are coupled through the three-species food-chain model. When only a topological analysis is in view, it does not matter too much which species is counted. But this usually requires thousands of observations (e.g. 30 000). One way of overcoming this limitation is to obtain a global model from a much smaller data set (e.g. less than 1000) and then use the model in the analysis. Nevertheless, when a global model is attempted, the choice of the species to count becomes very crucial. In particular, we have shown that if the generalist predator is counted, there is no hope for obtaining a model using a global modeling technique although in practice this might be the easiest species to count. This is quantified by observability indices which clearly indicate that the observability of the dynamics underlying this realistic three-species food chain is significantly less for the generalist predator located at the end of the food chain than for the other species. The best model has been obtained for the specialist predator which is located in the middle of the food chain. It should be noted that this specialist predator has direct interaction with the other two species involved in this food chain. It can be conjectured that the greater the interaction of a species, the better the observability of the dynamics. It is therefore important to choose properly the species to count for investigating a food chain to have a chance to recover the dynamics of the whole food chain, i.e. the dynamics of the non-counted species.

Acknowledgements

This work was partially supported by CNRS (France) and CNPq (Brazil).

References

[1] Packard NH, Crutchfield JD, Farmer JD, Shaw RS. Geometry from a time series. Phys Rev Lett 1980;45:712–16.
[2] Upadhyay RK, Jyengar SRK, Rai V. Chaos: an ecological reality? Int J Bifurc Chaos 1998;8(6):1325–33.
[3] Letellier C, Aziz-Alaoui MA. Analysis of the dynamics of a realistic ecological model. Chaos, Solitons & Fractals 2002;13:95–107.
[4] Aziz-Alaoui MA. Stability and nonlinear dynamics in a realistic three species food chain. SIAM J Math [submitted].
[5] Hastings A, Powell T. Chaos in three-species food chain. Ecology 1991;72:896–903.
[6] Rinaldi SR, Muratori S, Kuznetsov Y. Multiple attractors, catastrophes and chaos in seasonally perturbed predator prey communities. Bull Math Biol 1993;55:15–35.
[7] Klebanoff A, Hastings A. Chaos in three species food-chain. J Math Biol 1994;32:427–51.
[8] McCann K, Yodzis P. Biological conditions for chaos in a three-species food chain. Ecology 1994;75:561–4.
[9] Kuznetsov YA, Rinaldi S. Remarks on food chain dynamics. Math Biosci 1996;134:1–33.
[10] Aguirre LA, Billings SA. Retrieving dynamical invariants from chaotic data using NARMAX models. Int J Bifurc Chaos 1994;5(2):449–74.
[11] Letellier C, Maquet J, Le Sceller L, Gouesbet G, Aguirre LA. On the non-equivalence of observables in phase space reconstructions from recorded time series. J Phys A 1998;31:7913–27.
[12] Hamshi IL, Hamson L, Henttonen H. Specialist predators, generalist predators and the microtine rodent cycle. J Anim Ecol 1991;60:353–67.
[13] Jorgensen SE. Handbook of environmental data and ecological parameters. Oxford: Pergamon Press; 1979, p. 142–216.
[14] Gibson JF, Farmer JD, Casdagli M, Eubank S. An analytic approach to practical state space reconstruction. Physica D 1992;57:1–30.

[15] Takens F. Detecting strange attractors in turbulence. In: Rand DA, Young LS, editors. Dynamical systems and turbulence, Warwick 1980. Lecture notes in mathematics, vol. 898. New York: Springer; 1981. p. 366–81.
[16] Grassberger P, Proccacia I. Measuring the strangeness of strange attractors. Physica D 1983;9:189–208.
[17] Abarbanel HDI, Brown R, Sidorowich JJ, Tsimring LS. The analysis of observed chaotic data in physical systems. Rev Mod Phys 1993;65(4):1331–88.
[18] Cao L. Practical method for determing the minimum embedding dimension of a scalar time series. Physica D 1997;110(1–2):43–52.
[19] Scaffer WM, Kot M. Do strange attractors govern ecological systems? Ecology 1985;66:93–106.
[20] May RM. Chaos in natural populations. Proc R Soc London Ser A 1987;27:419–28.
[21] Gilmore R. Topological analysis of chaotic dynamical systems. Rev Mod Phys 1998;70(4):1455–529.
[22] Birman JS, Williams RF. Knotted periodic orbits in dynamical systems I: Lorentz's equations. Topology 1983;22(1):47–82.
[23] Melvin P, Tufillaro NB. Templates and framed braids. Phys Rev A 1991;44(6):3419–22.
[24] Mindlin GB, Solari HG, Natiello MA, Gilmore R, Hou XJ. Topological analysis of chaotic time series data from the Belousov–Zhabotinski reaction. J Nonlinear Sci 1991;1:147–73.
[25] Letellier C, Dutertre P, Maheu B. Unstable periodic orbits and templates of the Rössler system: toward a systematic topological characterization. Chaos, Solitons & Fractals 1995;5(1):271–82.
[26] Leontaritis IJ, Billings SA. Input–output parametric models for nonlinear systems part II: Stochastic nonlinear systems. Int J Control 1985;41(2):329–44.
[27] Gouesbet G, Maquet J. Construction of phenomenological models from numerical scalar time series. Physica D 1992;58:202–15.
[28] Aguirre LA. Controllability and observability of linear systems: some noninvariant aspects. IEEE Trans Educ 1995;38:33–9.
[29] Kailath T. Linear systems. Englewood Cliffs (NJ): Prentice-Hall; 1980.
[30] Chen S, Billings SA, Luo W. Orthogonal least squares methods and their application to nonlinear system identification. Int J Control 1989;50(5):1873–96.
[31] Aguirre LA. A nonlinear correlation function for selecting the delay time in dynamical reconstructions. Phys Lett A 1995;203 (2–3):88–94.
[32] Gouesbet G, Letellier C. Global vector field reconstruction by using a multivariate polynomial L_2 approximation on nets. Phys Rev E 1994;49(6):4955–72.
[33] Le Sceller L, Letellier C, Gouesbet G. Global vector field reconstruction taking into account a control parameter. Phys Lett A 1996;211:211–6.