Applied Mathematics and Computation 436 (2023) 127498

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Estimating Lyapunov exponents on a noisy environment by global and local Jacobian indirect algorithms

Lorenzo Escot[a], Julio E. Sandubete[a,b,*]

[a] Faculty of Statistical Studies, Complutense University of Madrid, Av. Puerta de Hierro 1, Madrid 28040, Spain
[b] Computing and Artificial Intelligence Lab, Camilo José Cela University, c/ Castillo de Alarcón, 49, Villanueva de la Cañada 28691, Spain

A R T I C L E I N F O

Article history:
Received 1 May 2022
Revised 4 August 2022
Accepted 17 August 2022

Keywords:
Chaotic time series
Lyapunov exponents
Jacobian indirect methods
Global and local neural net models
Local polynomial kernel models
Local neural net kernel models

A B S T R A C T

Most of the existing methods and techniques for the detection of chaotic behaviour from empirical time series try to quantify the well-known sensitivity to initial conditions through the estimation of the so-called Lyapunov exponents corresponding to the data generating system, even if this system is unknown. Some of these methods are designed to operate in noise-free environments, such as those methods that directly quantify the separation rate of two initially close trajectories. As an alternative, this paper provides two nonlinear indirect regression methods for estimating the Lyapunov exponents on a noisy environment. We extend the global Jacobian method, by using local polynomial kernel regressions and local neural net kernel models. We apply such methods to several noise-contaminated time series coming from different data generating processes. The results show that in general, the Jacobian indirect methods provide better results than the traditional direct methods for both clean and noisy time series. Moreover, the local Jacobian indirect methods provide more robust and accurate fit than the global ones, with the methods using local networks obtaining more accurate results than those using local polynomials.

1. Introduction

Chaos theory has been considered as the third greatest discovery on science after relativity and quantum mechanics, attracting ever increasing attention of many scientists from diverse disciplines, see e.g. [1–3]. Growing interest in chaotic approach has been propitiated, in part, because the solutions from linear deterministic systems, those traditionally used to model dynamic phenomena in the social and natural sciences, are perfectly regular, ordered and periodic. The linear deterministic models are, therefore, unable to reproduce the observed complexity, except by adding purely random behaviour. With the Chaos Theory, an alternative explanation for the aperiodic, irregular and seemingly random evolution is available but coming from a purely (nonlinear) deterministic system.

Particularly in statistics there is an increasing number of publications who are interested on chaos contributing to the development of statistical tools to deal with such systems., see e.g., [4–10]. The main challenge faced by those statistical tools is, precisely, the detection of chaotic behaviours in an irregular time series. That is, to test whether the irregular behaviour of a time series has its origin in deterministic chaotic behaviour, at least in part, or whether, on the contrary, the origin of the irregularity is a purely random stochastic processes.

* Corresponding author.
E-mail address: jsandube@ucm.es (J.E. Sandubete).

Several techniques have been developed to test for chaos, even when the data generating system is unknown and only an observed time-series is available see e.g., [11–13]. Most of these statistical tests try to quantify the initial-value sensitive property of chaotic systems estimating the *Lyapunov exponents* (λ) by using the Oseledec's multiplicative ergodic theorem [14]. Generally speaking this exponent measures how fast a perturbation in a point moves down the trajectory in a finite number of steps. For an in-depth discussion about the different definitions of Lyapunov exponents see some of the pioneering work in this field e.g., [15–19] or more up-to-date such as [20–23]. In this paper we have considered the definition proposed by Alligood et al. [24]. It can be defined as follows. Let $X_t = F(X_{t-1})$ be a difference equation where $F:\mathbb{R}^k \rightarrow \mathbb{R}^k$ for $t = 1, 2, 3, \ldots, T$. For this k-dimensional system there will be k Lyapunov exponents which are given by,

$$\begin{aligned}\lambda_i(X_0) &= \lim_{T\to\infty}\frac{1}{T}\{\log(|DF(X_T)|\cdot\ldots\cdot|DF(X_0)|)\}\\ &= \lim_{T\to\infty}\frac{1}{T}\left\{\log\left(\left|DF^T(X_0)\right|\right)\right\}\end{aligned} \tag{1}$$

where $DF^T(X_0)$ is the Jacobian evaluated along the trajectory $\{X_0, X_1, \ldots, X_T\}$ and $i = 1, 2, \ldots, k$. These exponents indicate the average rate of divergence or convergence of two nearby orbits starting close to the point X_0. It is important to note that theoretically proving this fact in the general case is really difficult, since Eq. (1) only considers linearisation and does not take into account e.g., Perron effects, although in our case, as it is a purely empirical study, it is not affected, for a discussion of this issue see [25]. Testing the hypothesis of chaotic behaviour in dissipative systems is equivalent to testing the hypothesis that at least one of those Lyapunov exponent is positive following [24]. A system will be dissipative if it shrinks or reduces asymptotically to a compact set under the action of the system over time. That is, it converges over time towards attractor sets, see [26]. For a discussion about definitions of a dissipative system in the sense of Levinson, see [27]. The null hypothesis for testing chaos will be

$$\begin{aligned}&H_0 : \hat{\lambda}_i > 0 \quad (\textit{Chaotic behaviour})\\ &\qquad\qquad H_1 : \hat{\lambda}_i \leqslant 0\end{aligned} \tag{2}$$

for some $i = 1, 2, 3, \ldots, k$. Reject the null hypothesis $H_0(\hat{\lambda}_i > 0)$ means lack of chaotic behaviour. When dynamic systems are known, we can directly calculate the full spectrum of the Lyapunov exponent value to check when the dynamic system is in a regime of chaotic behaviour applying directly Eq. (1). It should be noted that this situation is rarely possible in a real-world context because we do not know the system in most cases. Then it is very difficult to find any analytical solution to calculate the theoretical Lyapunov exponent. Therefore, we will focus on an empirical analysis of chaotic dynamical systems based on the study of observed time series data, which are usually a noise-contaminated signals of an unknown (deterministic or stochastic) system. The main objective is therefore to use an observed time series to test for chaotic behaviour (Eq. (2)) in the underlying unknown data-generating system.

There are two main methods in the literature that provide the estimated Lyapunov exponent from time-series data [28]. The first one, the so-called direct approach which directly measures the growth rate of the divergence between two trajectories with an infinitesimal difference in their initial conditions [26,29–33]. The main problem with this direct approach is that it only provides appropriate results when the series come from deterministic systems free of any noise. Any error or disturbance added to the series will be wrongly assigned to the existence of chaotic behavior.[1]

The second approach is the so called indirect or regression methods, that try to estimate the Jacobian in Eq. (1) by using the time series data [17]. Initially these jacobian indirect methods have focused on linear regressions [17,34–36]. Later other nonlinear neural nets regressions approach improved the estimation of Lyapunov exponents [28,37–40]. All of these indirect or Jacobian methods are global methods, in the sense that they use a regression model to estimate the Jacobian matrix of the system, and then this estimated Jacobian matrix is evaluated along the time series to obtain the Lyapunov exponents (Eq. (1)). These Jacobian methods solve the limitation of the direct approaches regarding noise. In this paper we are interested in analyzing and, where possible, improving the main advantage of these methods, namely their ability to estimate the exponents of an unknown underlying system with only an observer, a time series, an observer that in most cases could be contaminated by some kind of purely random noise (observation error)

In this sense, we have decided to extend the global indirect or Jacobian methods by proposing two new nonlinear local approaches based on polynomial kernel regressions and neural net kernel models. The main advantage of these local methods is that they provide a more accurate estimate of the underlying dynamics of the unknown system. Instead of estimating a single Jacobian matrix that is then evaluated over the entire observed orbit (global methods), local methods estimate a separate Jacobian matrix for each point of the orbit by using only the remainder locally nearby points in the phase space.

We have illustrated the robustness of the algorithms used for estimating the Lyapunov exponent from noise-contaminated time series. We are interested in comparing the results form direct methods, global nonlinear Jacobian indirect methods and the new two proposed local Jacobian methods. The results report that these local methods, specially local neural nets, provide us robust Lyapunov exponent estimators in the presence of measurement errors.

[1] Other approaches that do not rely on estimation the Lyapunov exponents to identify chaotic behaviours like the 0–1 test proposed by Gottwald & Melbourne [2004, 2005, 2008, 2009, 2016] have the same problem. They attribute any purely random observational error to the existence of deterministic chaos. see e.g., Hu et al. [2005] or Martinovič [2020].

The paper is structured in the following sections. Section 2 presents a discussion on the correct procedure for the reconstruction of the state space from noise-contaminated time series. It is also discussed some criteria in order to estimate the embedding parameters. Section 3 presents an overview about the main estimation methods of the Lyapunov exponent focusing on the Jacobian indirect methods based on the global neural net models, the local polynomial regressions and the neural net kernel models. Section 4 reports the main empirical results of this paper. Finally, Section 5 gives some concluding remarks.

2. How to get the state space reconstruction from noise-contaminated time series?

Since it was realised that even simple and deterministic dynamic systems can produce trajectories which look as a random process, there occurred some obvious questions: how can one distinguish chaos from randomness in real datasets?; how can we extract from time series relevant information about the unknown underlying generator system of them? In this section we are going to provide some of the key concepts which are required for answering those questions.

A dynamic system involves specifying a rule that governs the evolution of the process through time and the space in which the process takes values. If the rule that governs the evolution of the process is well understood it may be possible to write down the dynamic system more or less from first principles, without much need for observational data except perhaps to fix some parameters. Otherwise we may be much less clear about the underlying rules.

That is, we would have to deal with the following situations: (i) we might suspect that the system is governed by e.g., a set of difference equations, but not know what these are, even we may be uncertain what dynamic variables the equations involve; (ii) we may have available a quantity of experimental (or observational) data relating to how the system is observed to change through time. Then in that context we will have to assume that the true data-generating process is *unknown*. So we do not have the advantage of observing directly the functional form F that generate the dynamic associated with the state of the system X_T.

Instead of that, we have to consider an *observer function* $f : \mathbb{R}^{k+1} \to \mathbb{R}$ which turns the non-observed state variable of the dynamic system into observed data. This function f gives observations such as $x_t = f(X_t, \varepsilon_t)$ where the observable variable $x_t \in \mathbb{R}$, $t \in \mathbb{Z}^+$ (discrete-time) and ε_t (additive measurement error) is a sequence of independent and identically distributed (iid) random variables with mean equal to zero for $t = 1, 2, 3, \dots, n$. Hence we will consider that all information available is the noise-contaminated sequence $\{x_t\}_{t=1}^{n}$ as scalar strictly stationary univariate time series.

We have considered it appropriate to add a measurement noise term in the observer function because most real-world observed time series data are usually noise-contaminated signals. In this sense we want to know if as the measurement noise increases, the error committed in obtaining the estimator is amplified with the consequent inaccuracy or instead it is reduced.

Let us illustrate the effect of adding a small measurement noise to a well-known deterministic dynamic system as the Logistic map with a chaotic behaviour ($\mu = 4$). We have added to each time series data a normal multinomial error term denoted by $\varepsilon_t \sim N(0, s)$ with different variance values s. The data given in Fig. 1 provides the following remarks. First, as the measurement error increases, the dispersion of the huge amount of points that describe the attractor is amplified with the consequent inaccuracy. So a challenging question would be to quantify the amount of permissible measurement noise without smearing the qualitative characteristics of the attractor. Second, the noise amplification is uniform through the dynamics and it does not interact with the underlying generator system (it is independent on where we are). Third, when s (the variance of the error term) is greater than 0.1 the measurement noise can lead to confusion between a chaotic deterministic system and a purely stochastic one as you can see. In such cases the interplay between order and disorder, determinism and stochasticity, stability and instability or predictability and unpredictability is undergoing profound changes with very exciting implications on nonlinear time series analysis and chaos theory.

When the true data-generating process is unknown, it will not be possible to consider the true orbit of the dynamic system in the original state space. Instead of that, it will be possible to obtain a reconstruction of it that result equivalent in the dynamical and geometrical properties. Thus, this procedure ensures that the Lyapunov exponents previously described should be roughly the same in both the true and the reconstructed state spaces. This is an impressive fact obtained independently by Aeyels [41] and Takens [42] as point out [43]. It seems that Aeyels would have gotten the same type of results in his thesis in 1978 which was published later.

We have followed the framework proposed by Takens which extends to the field of dynamic systems the classic embedding theorem proposed by Whitney [44] in topology. In this connection, we would like to mention the following four interesting papers [45–48].

The embedding process enables testing of the chaos assumption (Eq. (2)) in the unknown original dynamical system. This fact allows scientists to construct models for complex and non-linear systems using just a single observable. Particularly, we have got the delayed-coordinate embedding vector considering the method of delayed-coordinates provided by Ruelle and Takens [49]. Let $\{x_t\}_{t=1}^{n}$ be a scalar strictly stationary univariate time series. We form a sequence of embedding vectors $\boldsymbol{x}_t^m$ by associating for each time period t a vector in a reconstructed state space $\mathbb{R}^m$, whose coordinates satisfy the following equation:

$$\boldsymbol{x}_t^m = \left(x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(m-2)\tau}, x_{t-(m-1)\tau}\right) \tag{3}$$

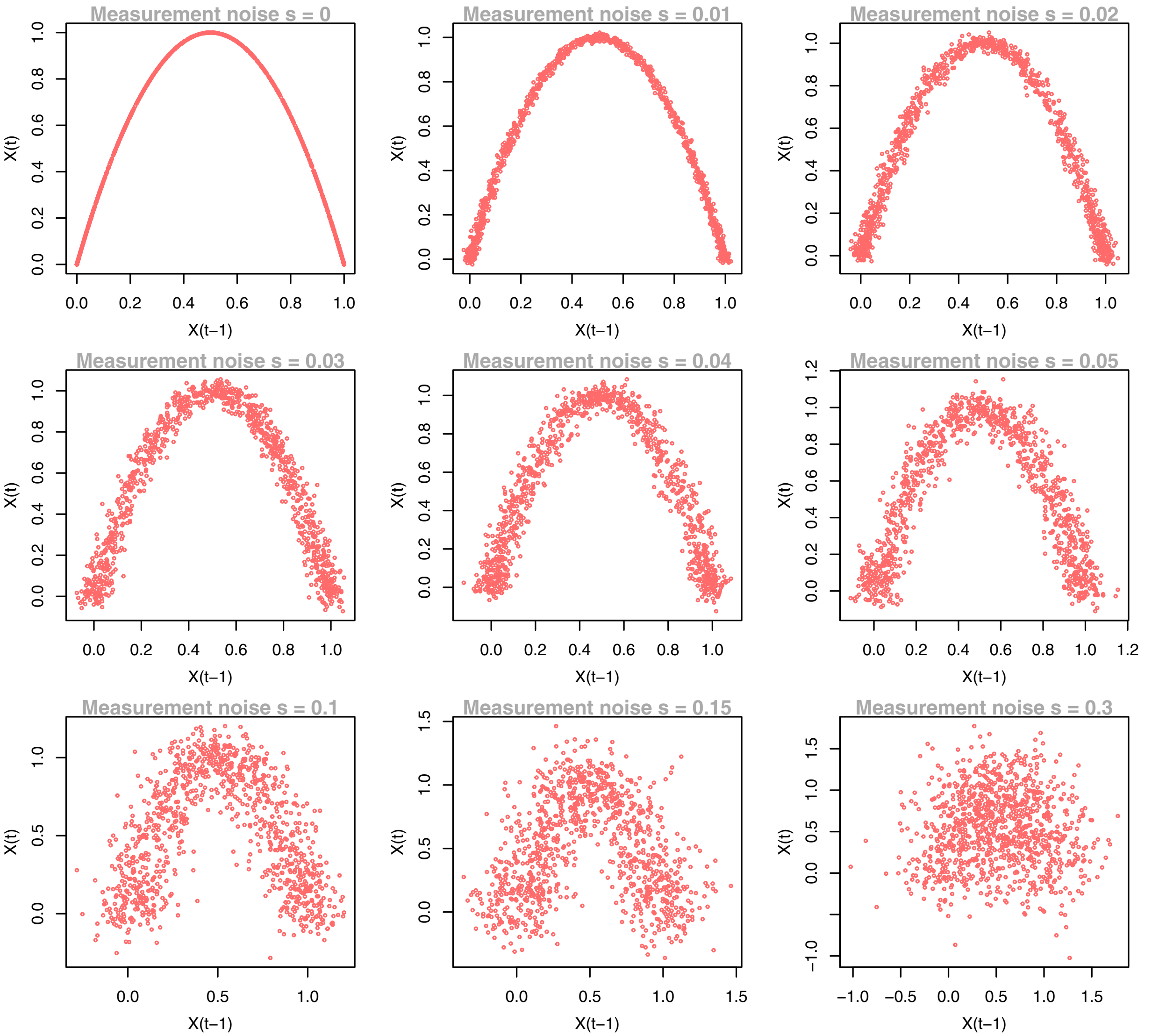

Fig. 1. Logistic map attractor adding a measurement noise with several variance values.

where τ is the reconstruction time-delay (or lag) and m is the embedding dimension. The number of embedding vectors will be equal to $n-(m-1)\tau$ and its length equal to the embedding dimension m. The relation between the time series $\{x_t\}_{t=1}^{n}$ and the underlying generator system is that $x_t = f(X_t, \varepsilon_t)$. This fact means that the reconstruction vectors of the time series are just the images under the delay embedding map of the successive points of the evolution X_t. So this map which is an embedding of $\mathcal{M} \subset \mathbb{R}^d$ into its image, sends the orbit of X_t to the sequence of reconstructions vectors $\left(x_t, x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-(m-2)\tau}, x_{t-(m-1)\tau}\right)$ of the time series $\{x_t\}_{t=1}^{n}$.

The underlying idea is to make copies of the single observable signal with uniform-time frequency ($t_i - t_{i-1} = t_s - t_{s-1}$ $\forall i \neq s$) and consider those delayed values as coordinates of a reconstructed state space retrieved from the time series. That is, despite the state being hidden from direct observation, the topology of the attractor that characterises the dynamic system can be preserved in the scalar univariate time serie when it is arranged into uniform delayed-coordinate embedding vectors.

A key point to create a suitable reconstruction of the state space is to fix a robust criteria in order to make the right choice of the embedding parameters. The implications of getting it wrong can be seen in the following Fig. 2. We have considered another well-known deterministic dynamic system as the Lorenz system with a chaotic behaviour ($\sigma = 10$, $\beta = 8/3$, $\rho = 28$) and different values for the time-delay τ from 1 to 1000.

For instance, on the one hand if we choose a time-delay τ equal to 1 we see practically a straight line. In this scenario, the dynamical and geometrical features are different between the real state space and the reconstructed ones. Hence, the Lyapunov exponents defined previously will not have the same value in both the true and the reconstructed state space. On

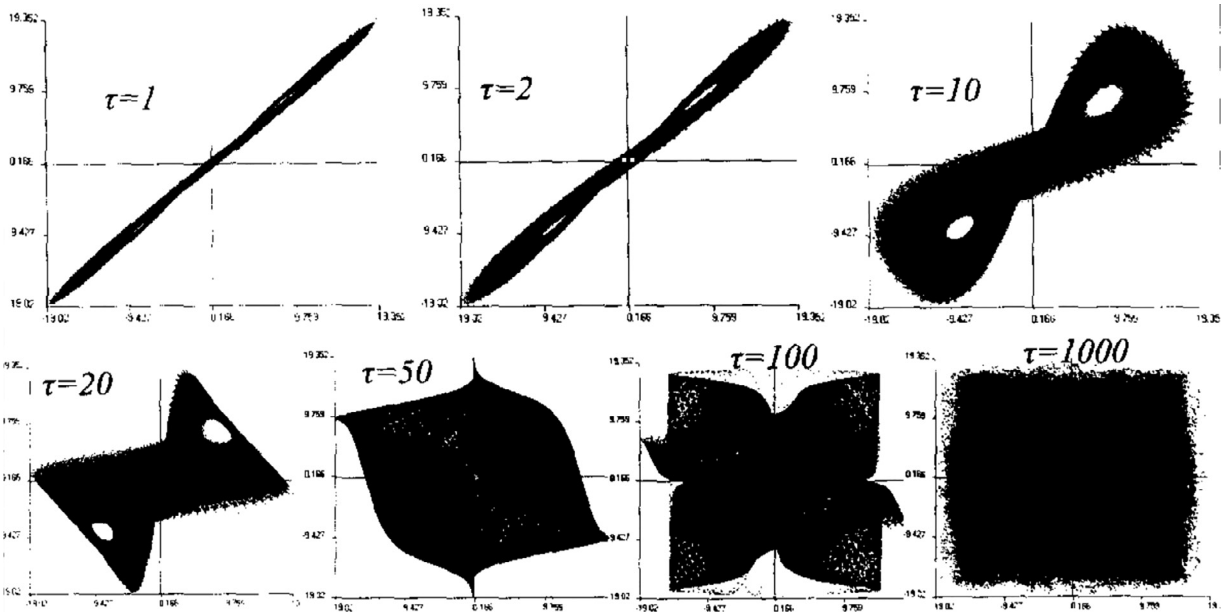

Fig. 2. Lorenz system's attractor reconstruction considering several time-delay values τ.

the other hand, if we choose a very large τ the state space becomes the same as the ones provided by a random behaviour. The same thing happens as in the previous case. Hence this is an important matter to be consider.

Researchers in this field usually determine them by two different alternatives: a heuristic approach that mostly rely on physical arguments and by a statistical approach. Under the heuristic approach regarding the choice of the time-delay τ although there are other criteria, see e.g., [19,50], $\tau = 1$ is commonly used following the prescription proposed by Takens [42]. Kennel et al. [51] proposed a method that is widely used to choose the embedding dimension m called false nearest neighbours. Another criteria widely used by the scientific community is to estimate the correlation dimension in increasing embeddings and to declare that the proper embedding has been found when that dimension saturates.

The key drawbacks of these heuristic methods are summarised as follows: (*i*) they do not take into account the results of any model fit; (*ii*) they lead to estimators whose properties are unknown or largely unexplored; (*iii*) they are not intrinsically statistical. The alternative proposed by the statistical approach solves those 3 disadvantages. The idea behind it is to select together (not independently) the embedding parameters τ and m which provide the best fit in the estimation of any quantitative measure e.g., the Lyapunov exponents taking into account some information criteria, see e.g., [52–54].

There are also model selection methods based on some information criteria that use cross-validation techniques which split the dataset into training, validation and test set. That is, we split the time series data into subsets, training the data on a subset and use the other subset to evaluate the model's performance. To reduce variability we should do multiple rounds of cross-validation with different subsets from the same dataset. Then we have to combine the validation results from these multiple rounds to come up with an estimate of the model's fit.

There are some cross validation techniques as e.g., leave one out cross validation (the most recommended cross-validation option in the reconstruction procedure), k-fold cross validation or stratified cross validation. In this paper we are going to use the statistical approach based on model selection procedures taking into account the Bayesian Information Criterion (BIC) instead of heuristic techniques, and the leave one out cross validation technique in order to choose the embedding parameter set.

3. How to get the Lyapunov exponents from delayed-coordinate embedding vectors?

Methods and techniques related to test the hypothesis of chaos (Eq. (2)) try to quantify the sensitivity to initial conditions estimating the so-called Lyapunov exponents. If one knows the data-generating process behind the time series the theoretical Lyapunov exponent can be calculated directly using its own definition, see Eq. (1). However we have assumed that the true dynamics of the system is unknown.

As we have mentioned in the introduction the first main method to estimate Lyapunov exponent from time series, is the so-called *direct* approach which directly measures the growth rate of the divergence between two nearby trajectories with an infinitesimal difference in their initial conditions [5,26,29–33]. The main disadvantage of these direct method are, first, that they only allows the estimation of largest Lyapunov exponents (not the full spectrum), and second, that it only provides appropriate results with large the series coming from deterministic systems free of any noise. They assign to chaos any divergence in two nearby orbits, even if purely random, caused, for example, by the measurement error itself.

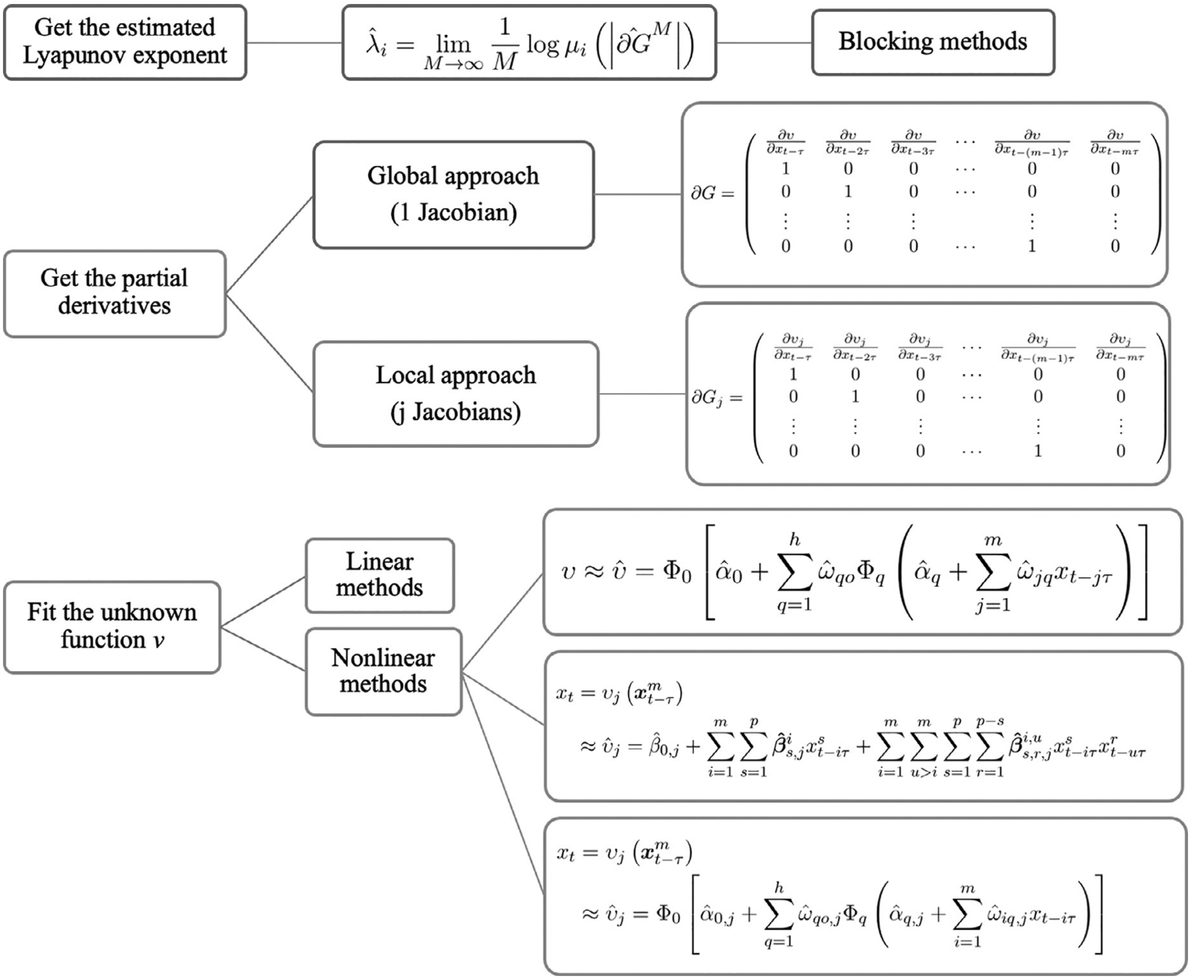

Fig. 3. Theoretical framework for estimating the Lyapunov exponents from time series by Jacobian indirect methods.

As an alternative to these direct methods, the *Jacobian or regression indirect method* proposed initially by Eckmann and Ruelle [17] allows the estimation of the full spectrum of Lyapunov exponents, and provides consistent estimators indeed when the deterministic time series is contaminated with (small) measurement noise, or the time series has a moderate sample sizes. Aditionally, the asymptotic distribution of those Jacobian estimators can be derivedShintani2004, allowing to make a formal test for chaos using (2).

The idea behind this Jacobian indirect approach method can be describes as follows. To estimate Lyapunov exponents we just have to evaluate the derivatives (Eq. (1)) along the full path because those derivatives will give us an indication of how two initial closed paths diverge. The two main issues are the followings: (*i*) how to compute the Jacobians and (*ii*) how to assess them over the trajectory. For the first question we distinguish between lineal, polynomial and neural net models. For the last question we distinguish between global and local (see 3). We are going to discuss the main features of those kind of indirect methods for estimating the Lyapunov exponents considering both a global and local regression setting in the next sections.

3.1. Estimating Lyapunov exponents by global Jacobian algorithms

Our starting point is a data generation process that is unknown a priori. In this sense, we have to consider an observer function that include a measurement noise term ε_t which generates observations as $x_t = f(X_t, \varepsilon_t)$ where ε_t is a sequence of independent and identically distributed random variables for $t = 1, 2, 3, \ldots, n$. Then, we form a sequence of delayed vectors by associating for each time period a vector in a reconstructed state space $\mathbb{R}^m$, whose coordinates satisfy the reconstruction equation $\boldsymbol{x}_t^m = \left(x_t, x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-(m-2)\tau}, x_{t-(m-1)\tau}\right)$.

Let assume following Gençay and Dechert [28] that there exist a a pseudo dynamic system $G : \mathbb{R}^m \to \mathbb{R}^m$ such that $\boldsymbol{x}_t^m = G\left(\boldsymbol{x}_{t-\tau}^m\right)$ where $\boldsymbol{x}_t^m$ are the uniform delayed-coordinate embedding vectors taking $t_i - t_{i-1} = t_s - t_{s-1}$ $\forall i \neq s$. The dynamic system G could be expressed as a matrix that depends on a function υ which, in turn, is a function of the delays,

$$\begin{pmatrix} x_t \\ x_{t-\tau} \\ \vdots \\ x_{t-(m-2)\tau} \\ x_{t-(m-1)\tau} \end{pmatrix} = G \begin{pmatrix} x_{t-\tau} \\ x_{t-2\tau} \\ \vdots \\ x_{t-(m-1)\tau} \\ x_{t-m\tau} \end{pmatrix}$$

$$\begin{pmatrix} x_t \\ x_{t-\tau} \\ \vdots \\ x_{t-(m-2)\tau} \\ x_{t-(m-1)\tau} \end{pmatrix} = \begin{pmatrix} \upsilon\left(x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-(m-2)\tau}, x_{t-(m-1)\tau}, x_{t-m\tau}\right) \\ x_{t-\tau} \\ \vdots \\ x_{t-(m-2)\tau} \\ x_{t-(m-1)\tau} \end{pmatrix} \tag{4}$$

The global Jacobian corresponding to the reconstructed dynamic system G will be obtained as follows,

$$\partial G = \begin{pmatrix} \frac{\partial \upsilon}{\partial x_{t-\tau}} & \frac{\partial \upsilon}{\partial x_{t-2\tau}} & \frac{\partial \upsilon}{\partial x_{t-3\tau}} & \cdots & \frac{\partial \upsilon}{\partial x_{t-(m-1)\tau}} & \frac{\partial \upsilon}{\partial x_{t-m\tau}} \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{5}$$

Then the estimation of the Lyapunov exponent by the global Jacobian indirect method is reduced to the estimation of the unknown nonlinear function $\upsilon : \mathbb{R}^m \to \mathbb{R}$. This is the key. This means that the υ function that contains the relation between x_t and its past must be estimated. The different approaches involved in the global indirect method diverge in the algorithm used for estimating the υ function in the Jacobian Eq. (5). Particularly, we have set our focus on a non-linear method that try to estimate the underlying dynamic system without imposing the restriction of linearity considering neural net models.

3.1.1. Approximating the unknown function υ through a global neural net model

Following McCaffrey et al. [55], Nychka et al. [37], Gençay and Dechert [28] and Shintani and Linton [39,40], the *global neural net estimator* can be obtained by approximating the unknown nonlinear function υ in the Jacobian (Eq. (5)) through a feedforward single hidden layer network with a single output as follows,

$$\upsilon \approx \hat{\upsilon} = \Phi_0\left[\hat{\alpha}_0 + \sum_{q=1}^{h} \hat{\omega}_{qo}\Phi_q\left(\hat{\alpha}_q + \sum_{j=1}^{m} \hat{\omega}_{jq}x_{t-j\tau}\right)\right] \tag{6}$$

where $\hat{\omega}_{jq}$ are the estimated layers connection weights from hidden layer to output, $\hat{\alpha}_q$ is the estimated network bias from hidden layer, m is the embedding dimension, Φ_q is the transfer function. We have used the logistic function. The number of nodes in the single hidden layer is denoted by h, $\hat{\omega}_{q0}$ are the expected layer connection weights from the input to the hidden layer, $\hat{\alpha}_0$ is the estimated bias of the neural net from the input and $\Phi_0 \in I$. The issue of estimating the parameters can be reduced to a least squares problem where the quantity to be minimised is given by

$$\sum_{t=1}^{n}\left(x_t - \left[\alpha_0 + \sum_{q=1}^{h} \omega_{qo}\Phi_q\left(\alpha_q + \sum_{j=1}^{m} \omega_{jq}x_{t-j\tau}\right)\right]\right)^2$$

Then we have obtained the partial derivates of the Jacobian in Eq. (5) applying the chain rule to Eq. (6) as follows,

$$\frac{\partial \hat{\upsilon}}{\partial x_{t-j\tau}} = \Phi'_0(z_0)\sum_{q=1}^{h} \hat{\omega}_{q0}\Phi'_0(z_q)\hat{\omega}_{jq} \tag{7}$$

where

$$z_0 = \hat{\alpha}_0 + \sum_{q=1}^{h} \hat{\omega}_{q0}\Phi_q(z_q), \quad z_q = \hat{\alpha}_q + \sum_{j=1}^{m} \hat{\omega}_{jq}x_{t-j\tau}$$

The estimated partial derivates are given by (see Fig. 3)

$$\partial \hat{G} = \begin{pmatrix} \frac{\partial \hat{\upsilon}}{\partial x_{t-\tau}} & \frac{\partial \hat{\upsilon}}{\partial x_{t-2\tau}} & \frac{\partial \hat{\upsilon}}{\partial x_{t-3\tau}} & \cdots & \frac{\partial \hat{\upsilon}}{\partial x_{t-(m-1)\tau}} & \frac{\partial \hat{\upsilon}}{\partial x_{t-m\tau}} \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{8}$$

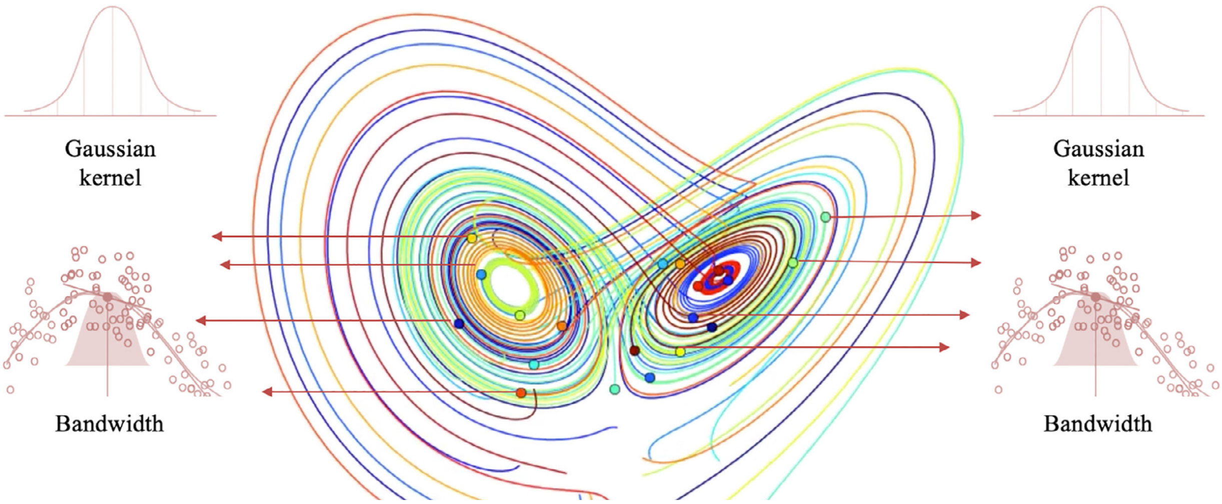

Fig. 4. Diagram about the kernel-type regression: the data point farther away from an arbitrary point receive less weight than the data point closer to it considering a certain kernel function and a bandwidth.

The results proposed by Hornik et al. [56] have enabled us to consider neural net models with just one single hidden layer. The number of hidden units is established using statistical methods according to model selection criteria, as it appears in the results of this document. As we noted in [57]: *the main reason for using neural net models is not to look for the best predictive model but to estimate a model that captures the nonlinear time dependence well enough and, additionally, allows us to obtain in an analytical way (instead of numerical) the Jacobian functional of the unknown data-generating process* (Eq. (5)). *The estimation of this Jacobian will allow us to contrast the hypothesis of chaos* (Eq. (2)).

3.2. Estimating Lyapunov exponents by local Jacobian algorithms

Having discussed the global Jacobian indirect approach in the previous section we will now focus on greater detail into the local context of the Jacobian indirect methods. Particularly, the local Jacobian approach has the following advantages over the global Jacobian methods: (*i*) it is superior in the evaluation of the system's derivatives at each point along the whole trajectory. This fact is relevant in order to provide better fits to point-wise estimates in this context because by definition when estimating the Lyapunov exponents through the indirect methods we have to make an evaluation of the Jacobian at each point along the whole trajectory; (*ii*) it captures the nonlinear time dependence well enough considering the condition of local nonlinearity.

In this section we are going to describe the key features of the local Jacobian procedure for estimating the Lyapunov exponents from noise-contaminated time series. Remember that to get the estimated Lyapunov exponents the key is to estimate the pseudo dynamic system denoted by G. In the global context we estimate just a *single* Jacobian (Eq. (8)) and then we evaluate it along the whole trajectory. Now in a local context, we have to estimate a Jacobian for *each point* along the full path. We mean one function υ for each point and how do we do that? We have extended the results proposed by Gençay and Dechert [28] considering a local context. Let us illustrate it.

By definition when estimating the Lyapunov exponents through the indirect methods we have to make an evaluation of the Jacobian at each point along the whole trajectory. In this sense, the local Jacobian indirect method seems to be suitable for estimating the Lyapunov exponents because this approach is based on essentially the weighted least squares regression technique (point-wise estimators). In this kernel-type regression the data points farther away from an arbitrary point receive less weight than data points closer to it. As with density estimation, the weighting of the points is done through a kernel function, see Fig. 4. The choice of which kernel K to use has a relatively small effect on the result of a kernel regression. In our case, we have considered the Gaussian kernel.

However, the choice of the bandwidth $\hbar$ that controls the weight (distance or neighbourhood size) around the point has a larger effect on results. It also remains true that bandwidths $\hbar$ that are too large produce over-smoothed estimates while if it is too small the estimates appear excessively bumpy or under-smoothed. In this sense, there is a vast set of publications about methods for bandwidth choice in kernel regression estimation, for a review see e.g., [58–62].

As far as this paper is concerned the bandwidth $\hbar$ (together with all other parameters) will be determined by the statistical approach based on model selection procedures taking into account the Bayesian information criterion and leave one out cross validation technique. The procedure can be summarized as follows. For each setting the optimal bandwidth is calculated. Then, the weights for all points are calculated considering the optimal bandwidth, the well-known Gaussian kernel and the leave one out cross validation technique. These weights are used in the nonlinear weighted least squares regressions. After estimating the coefficients, the residuals are taken out and the BIC criterion is implemented as follows,

$$BIC = \log(RSS) + \frac{\log(n)}{n}[1 + coef(m+2)]$$

with RSS being the residual sum of squares, m the embedding dimension, n the total number of observations, and $coef$ the coefficients included in each weighted least squares regression, as follows.

Now let us explain how we have got the Jacobians considering the local approach. As in the global case the key is the function υ. We assume that for each point x_j there are jth functions $G_j : \mathbb{R}^m \to \mathbb{R}^m$ such that the reconstructed dynamic system can be defined locally around a point x_j as $\boldsymbol{x}_t^m = G_j\left(\boldsymbol{x}_{t-\tau}^m\right)$ where $\boldsymbol{x}_t^m$ is defined by Eq. (3). The dynamic system G_j could be expressed around a point x_j as a matrix that depends on the functions υ_j which, in turn, are functions of the delays,

$$\begin{pmatrix} x_t \\ x_{t-\tau} \\ \vdots \\ x_{t-(m-2)\tau} \\ x_{t-(m-1)\tau} \end{pmatrix} = G_j \begin{pmatrix} x_{t-\tau} \\ x_{t-2\tau} \\ \vdots \\ x_{t-(m-1)\tau} \\ x_{t-m\tau} \end{pmatrix}$$

$$\begin{pmatrix} x_t \\ x_{t-\tau} \\ \vdots \\ x_{t-(m-2)\tau} \\ x_{t-(m-1)\tau} \end{pmatrix} = \begin{pmatrix} \upsilon_j\left(x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-(m-2)\tau}, x_{t-(m-1)\tau}, x_{t-m\tau}\right) \\ x_{t-\tau} \\ \vdots \\ x_{t-(m-2)\tau} \\ x_{t-(m-1)\tau} \end{pmatrix} \tag{9}$$

The local partial derivatives based on the Jacobians from the reconstructed dynamic system G_j around a point x_j will be obtained as follows,

$$\partial G_j = \begin{pmatrix} \frac{\partial \upsilon_j}{\partial x_{t-\tau}} & \frac{\partial \upsilon_j}{\partial x_{t-2\tau}} & \frac{\partial \upsilon_j}{\partial x_{t-3\tau}} & \cdots & \frac{\partial \upsilon_j}{\partial x_{t-(m-1)\tau}} & \frac{\partial \upsilon_j}{\partial x_{t-m\tau}} \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{10}$$

As you can see, we would have as many Jacobians as points along the whole trajectory considered in the local context. Instead, there would be only one in the global Jacobian (Eq. (5)). The main difference between both approaches can be summarised as follows. We have to estimate just a single model υ for all data in the global context. Then, the Jacobian ∂G is evaluated at each point x_j for $j = 1, 2, 3, \ldots, n$. In the local context, we have to make a local regression model υ_j around each point x_j (with decreasing weights as the point moves away from x_j).

Then, the Jacobian ∂G_j is evaluated exclusively at point x_j, and this procedure is done for all points along the whole trajectory considered. So the estimation of the Lyapunov exponents by the local approach is reduced to the estimation of the unknown nonlinear functions $\upsilon_j : \mathbb{R}^m \to \mathbb{R}$. This is the key again as in the global case. This means that we have to compute those functions υ_j that collect the relationship between x_t and its past locally.

The different approaches that compose the local indirect methods differ in the algorithm used for the estimation of the functions υ_j in the Jacobians (Eq. (10)). Traditionally the main contributions proposed by the scientific community regarding the local Jacobian indirect methods have focused on local linear approaches. The local linear estimator of the Lyapunov exponent was first proposed by Lu and Smith [63], Lu [64], and then revisited by Whang and Linton [65] and Shintani and Linton [39].

The main drawbacks of these local linear estimators of the Lyapunov exponent are the following: (i) the estimated partial derivatives will be a constant for each embedding dimension considered but it will depend at each point x_j; (ii) approximate linearly the functions υ_j locally at each point x_j through the result provided by linear weighted least squares regressions may not be the best approach since it should not be forgotten that chaotic behaviour is a phenomenon in nonlinear dynamic systems. It does not exist in linear systems. Due to these disadvantages we are going to extend those results by proposing two novel alternatives that differ in the algorithm used for the estimation of those functions υ_j in the local Jacobian. Particularly, we are going to introduce the following two nonlinear local approaches based on polynomial kernel regressions and neural net kernel models. Let us explain in detail these local non-parametric techniques.

3.2.1. Approximating the unknown function υ through a local polynomial kernel regression

Firstly, we have applied a class of kernel-type regression estimators called *polynomial kernel* approach based on the local Jacobian indirect methods. In this sense, we have considered a polynomial of order p greater than one generalising the *m-step Lyapunov-like index* proposed by Yao and Tong [66] for an embedding dimension $m \geqslant 1$. The main reason for using polynomial kernel regressions is not to look for the best predictive model but to estimate a model that captures the nonlinear time dependence well enough and, additionally, allows us to obtain in an analytical way (instead of numerical) the local Jacobians from the unknown data-generating process provided by Eq. (10). We have developed our own algorithm as follows.

Let consider $x_t = \upsilon_j\left(x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-(m-2)\tau}, x_{t-(m-1)\tau}, x_{t-m\tau}\right) = \upsilon_j\left(\boldsymbol{x}^m_{t-\tau}\right)$ as a sequence of delayed values obtained from the reconstructed dynamic system defined locally for $m \geqslant 1$. The *polynomial kernel estimator* can be obtained by approximating nonlinearly the functions υ_j locally around a point x_j through the result provided by jth nonlinear weighted least squares regressions as follows (considering the cross-product terms)

$$x_t = \upsilon_j\left(\boldsymbol{x}^m_{t-\tau}\right)$$

$$\approx \hat{\upsilon}_j = \hat{\beta}_{0,j} + \sum_{i=1}^{m}\sum_{s=1}^{p}\hat{\boldsymbol{\beta}}^i_{s,j}x^s_{t-i\tau} + \sum_{i=1}^{m}\sum_{u>i}^{m}\sum_{s=1}^{p}\sum_{r=1}^{p-s}\hat{\boldsymbol{\beta}}^{i,u}_{s,r,j}x^s_{t-i\tau}x^r_{t-u\tau} \tag{11}$$

where the betas $\hat{\boldsymbol{\beta}}$ are the estimated coefficients for $i = 1, 2, \ldots, m;\ s = 1, 2, \ldots, p;\ j = 1, 2, \ldots, n-(m-1)\tau$. Note that the issue of parameter estimation is reduced again to a nonlinear weighted least squares problem in which the quantity to be minimized for each point x_j is given by

$$\min_{\hat{\boldsymbol{\beta}}}\sum_{t=1}^{n-(m-1)\tau}\left[\left(x_t - \hat{\upsilon}_j\right)^2 \cdot K\left(\frac{\boldsymbol{x}^m_{t-\tau} - \boldsymbol{x}^m_{j-\tau}}{\hbar}\right)\right]$$

where $K(\cdot)$ is the kernel function and $\hbar$ denotes the bandwidth. We have considered the Gaussian kernel expressed in this context by

$$K\left(\frac{\boldsymbol{x}^m_{t-\tau} - \boldsymbol{x}^m_{j-\tau}}{\hbar}\right) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x^m_{t-\tau}-x^m_{j-\tau}}{\hbar}\right)^2}$$

The bandwidth $\hbar$ will be determined by the statistical approach based on model selection procedures taking into account the Bayesian information criterion and the leave one out cross-validation technique as mentioned earlier. We have obtained the partial derivatives provided by the local Jacobian defined in Eq. (10) from the estimated functions $\hat{\upsilon}_j$ defined locally by Eq. (11) as follows (considering the cross-product terms)

$$\frac{\partial\hat{\upsilon}_j}{\partial x_{t-i\tau}} = \hat{\boldsymbol{\beta}}^i_{1,j} + \sum_{s=1}^{p}s\cdot\hat{\boldsymbol{\beta}}^i_{s,j}x^{s-1}_{t-i\tau} + \sum_{u>i}^{m}\sum_{s=1}^{p}\sum_{r=1}^{p-s}s\cdot\hat{\boldsymbol{\beta}}^{i,u}_{s,r,j}x^{s-1}_{t-i\tau}x^r_{t-u\tau} \tag{12}$$

where $i = 1, 2, \ldots, m;\ s = 1, 2, \ldots, p;\ j = 1, 2, \ldots, n-(m-1)\tau$. If $p = 1$ we refer to the linear case.

The estimated partial derivates $\partial\hat{G}_j$ based on the local Jacobian defined in Eq. (10) for $m \geqslant 1$ are given by (see Fig. 3)

$$\partial\hat{G}_j = \begin{pmatrix} \frac{\partial\hat{\upsilon}_j}{\partial x_{t-\tau}} & \frac{\partial\hat{\upsilon}_j}{\partial x_{t-2\tau}} & \frac{\partial\hat{\upsilon}_j}{\partial x_{t-3\tau}} & \cdots & \frac{\partial\hat{\upsilon}_j}{\partial x_{t-(m-1)\tau}} & \frac{\partial\hat{\upsilon}_j}{\partial x_{t-m\tau}} \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{13}$$

Note that if we would consider the one-dimensional linear case ($m = 1,\ p = 1$) the estimated Lyapunov exponent can be got directly as follows

$$\hat{\lambda} = \lim_{n\to\infty}\frac{1}{n}\log\mu\left(\left|\partial\hat{G}\right|\right) = \lim_{n\to\infty}\frac{1}{n}\log\mu\left(\left|\hat{\beta}_{1,j}\right|\right) = \lim_{n\to\infty}\frac{1}{n}\log\hat{\beta}_{1,j} \tag{14}$$

where μ is the eigenvalue provided by the estimated partial derivates $\partial\hat{G}_j = \left(\frac{\partial\hat{\upsilon}_j}{\partial x_{t-\tau}}\right) = \hat{\beta}_{1,j}$ for $j = 1, 2, \ldots, n$ (full sample size). There will be just one Lyapunov exponent because the embedding dimension $m = 1$. Note that the local linear estimator of the Lyapunov exponent is reduced to the mean-log (mean of the logarithm) of the estimated coefficients $\hat{\beta}_{1,j}$ due to the eigenvalue of a number is the number itself. This result is exactly the same as the local linear estimator called *m-step Lyapunov-like index* proposed by Yao and Tong [66]. As we are not considering just an one-dimensional or linear context we have to evaluate the Jacobians at each points along the whole trajectory and get the Lyapunov exponent estimator as explained below. Before that, let us focus on the second local estimator.

3.2.2. Approximating the unknown function υ through a local neural net kernel model

In this section we are going to propose another novel local nonlinear approach based on neural net kernel models. As far as we know, the local neural net kernel estimator of the Lyapunov exponent has not been proposed by anyone so far. This approach has some significant advantages over the global neural net case seen before: (i) it improves the fit over the global neural net estimator as we will see later on; (ii) it is superior to the global neural net estimator in the evaluation of the partial derivatives at each point along the whole trajectory for the reasons described above; (*iii*) it estimates the functions υ_j considering the condition of local nonlinearity. Note that the idea behind the local neural net procedure is similar to the case described in the previous section because both are based on nonlinear kernel regressions, see Cheng et al. [67].

The neural net kernel approach for the estimation of the functions υ_j in the local Jacobian (Eq. (10)) is based essentially on point-wise estimates through a local standard feed-forward neural network with as few as one hidden layer instead of a local polynomial, where the data points farther away from an arbitrary point x_j receive less weight than data points closer to x_j as we mentioned before for the local polynomial. We have developed our own algorithm as follows.

Let consider a sequence $x_t = \upsilon_j\left(x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-(m-2)\tau}, x_{t-(m-1)\tau}, x_{t-m\tau}\right) = \upsilon_j\left(\boldsymbol{x}_{t-\tau}^m\right)$ of delayed values obtained from the reconstructed dynamic system defined locally for $m \geqslant 1$. The *neural net kernel estimator* can be obtained by approximating *nonlinearly* the functions υ_j locally around a point x_j through the result provided by jth feed-forward single hidden layer neural networks with a single output by

$$\begin{aligned} x_t &= \upsilon_j\left(\boldsymbol{x}_{t-\tau}^m\right) \\ &\approx \hat{\upsilon}_j = \Phi_0\left[\hat{\alpha}_{0,j} + \sum_{q=1}^{h} \hat{\omega}_{q0,j}\Phi_q\left(\hat{\alpha}_{q,j} + \sum_{i=1}^{m} \hat{\omega}_{iq,j}x_{t-i\tau}\right)\right] \end{aligned} \tag{15}$$

where $\hat{\omega}_{iq,j}$ are the estimated layers connection weights from hidden layer to output at each point x_j, m is the embedding dimension, $\hat{\alpha}_{q,j}$ are the estimated network coefficients from hidden layer at each point x_j, Φ_q is the transfer function which in our case is the logistic function, $\hat{\omega}_{q0,j}$ are the estimated layers connection weights from input to hidden layer at each point x_j, the number of neurones (or nodes) in the single hidden layer is denoted by h, $\hat{\alpha}_{0,j}$ is the estimated network coefficients from input at each point x_j and $\Phi_0 \in I$. The issue of parameter estimation is reduced to a nonlinear least squares problem in which the quantity to be minimized for each point x_j is given by

$$\sum_{t=1}^{n-(m-1)\tau}\left[\left(x_t - \left(\alpha_{0,j} + \sum_{q=1}^{h}\omega_{q0,j}\Phi_q\left(\alpha_{q,j} + \sum_{i=1}^{m}\omega_{iq,j}(x_{t-i\tau})\right)\right)\right)^2 \cdot K\left(\frac{\boldsymbol{x}_{t-\tau}^m - \boldsymbol{x}_{j-\tau}^m}{\hbar}\right)\right]$$

where $K(\cdot)$ is the kernel function and $\hbar$ denotes the bandwidth. We have considered again the Gaussian kernel expressed in this context by

$$K\left(\frac{\boldsymbol{x}_{t-\tau}^m - \boldsymbol{x}_{j-\tau}^m}{\hbar}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_{t-\tau}^m - x_{j-\tau}^m}{\hbar}\right)^2}$$

The bandwidth $\hbar$ will be determined by the procedure explained previously. We have obtained the partial derivatives provided by the local Jacobian defined in Eq. (10) around a point x_j applying the chain rule to Eq. (15) as follows,

$$\frac{\partial \hat{\upsilon}_j}{\partial x_{t-i\tau}} = \Phi'_0\left(z_{0,j}\right)\sum_{q=1}^{h}\hat{\omega}_{q0,j}\Phi'_0\left(z_{q,j}\right)\hat{\omega}_{iq,j} \tag{16}$$

where

$$z_{0,j} = \hat{\alpha}_{0,j} + \sum_{q=1}^{h}\hat{\omega}_{q0,j}\Phi_q\left(z_{q,j}\right), \quad z_{q,j} = \hat{\alpha}_{q,j} + \sum_{i=1}^{m}\hat{\omega}_{iq,j}x_{t-i\tau}$$

for $i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n-(m-1)\tau$ and the estimated partial derivates $\partial\hat{G}_j$ based on the local Jacobian defined in Eq. (10) around a point x_j for $m \geqslant 1$ are given by

$$\partial\hat{G}_j = \begin{pmatrix} \frac{\partial\hat{\upsilon}_j}{\partial x_{t-\tau}} & \frac{\partial\hat{\upsilon}_j}{\partial x_{t-2\tau}} & \frac{\partial\hat{\upsilon}_j}{\partial x_{t-3\tau}} & \cdots & \frac{\partial\hat{\upsilon}_j}{\partial x_{t-(m-1)\tau}} & \frac{\partial\hat{\upsilon}_j}{\partial x_{t-m\tau}} \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{17}$$

3.3. Estimating and making inference about the Lyapunov exponent

Now let us provide the right procedure to obtain a consistent estimator of the Lyapunov exponent defined by Eq. (1) considering the partial derivatives calculated analytically from the best-fitted global neural net model (Eq. (7)) or the two nonlinear local approaches seen before, polynomial kernel regression (Eq. (12)) and local neural net kernel model (Eq. (16)). The kth Lyapunov exponent estimator can be obtained as follows,

$$\hat{\lambda}_k = \lim_{M\to\infty}\frac{1}{M}\log\mu_k\left(\left|\partial\hat{G}^M\right|\right) \tag{18}$$

Table 1
Theoretical Lyapunov exponents values (λ_{th}) from some well-known chaotic dynamic systems.

Dynamic system	Equations	Parameters	λ_{th}
Logistic	$x_t = \mu x_{t-1}(1 - x_{t-1})$	$\mu = 4$	0.69314
Gauss	$x_t = e^{-\alpha x_{t-1}^2} + \beta$	$\alpha = 6.2, \beta = -0.5$	0.38367
Hénon	$x_t = 1 - a x_{t-1}^2 + y_{t-1}$	$a = 1.4$	0.41921
	$y_t = b x_{t-1}$	$b = 0.3$	-1.63479
Rössler	$\dot{x} = -y - z$	$a = 0.2$	0.07143
	$\dot{y} = x + ay$	$b = 0.2$	0.00000
	$\dot{z} = b + (x - c)z$	$c = 5.7$	-0.53943

where μ_k is the kth largest eigenvalue provided by the global jacobian $\partial\hat{G}^M = \partial\hat{G}(x_M) \cdot \partial\hat{G}(x_{M-1}) \cdot \ldots \cdot \partial\hat{G}(x_1)$ for $k = 1, 2, 3, \ldots, m$ where $\partial\hat{G}(\cdot)$ are the estimated partial derivatives calculated from the best-fit neural net model (Eq. (8)). In the local context μ_k would be the kth largest eigenvalue by the local jacobians $\partial\hat{G}_j^M = \partial\hat{G}_j(x_{j,M}) \cdot \partial\hat{G}_j(x_{j,M-1}) \cdot \ldots \cdot \partial\hat{G}_j(x_{j,1})$ around a point x_j where $\partial\hat{G}_j(\cdot)$ are the partial derivates estimated from the best-fitted polynomial or neural net kernel models (Eq. (13)), m is the embedding dimension and M is the block length.

The choice of the block length M is an important issue to guarantee the consistency of the estimation, for a review see e.g. [38,55], [65,68]. Note that we can make a distinction between the sample size n used to estimate the partial derivatives of the Jacobian in Eq. (8) (global case) or Eq. (13) (local context) and the block length M, defined in Eq. (18), which is the number of evaluation points used for estimating the kth Lyapunov exponents.

In this paper we have considered both the full sample ($M = n$) and three different methods of subsampling by blocks: non-overlapping, equally spaced and bootstrap blocking methods. It can be described in few words the main idea behind those methods as follows. We need to split the full sample into blocks, then we have to estimate a Lyapunov exponent for each block and finally we should take the mean or median of the estimated exponents for each block. The choice of the observations within every block is what differentiates each of the three blocking methods under consideration.

Finally, we have followed the ideas proposed by Shintani and Linton [40] regarding the asymptotic properties of the Lyapunov exponent estimator we have just seen. The asymptotic properties of the Lyapunov exponent estimator are *common* to all Jacobian indirect methods regardless of the approach used for estimating the function υ in the global setting or the functions υ_j in the local environment, see Shintani and Linton [39].

We particularly seek to know if the estimated Lyapunov exponents are (or are not) statistically significant higher than 0. In this sense, our objective will be to test the null hypothesis $H_0 : \hat{\lambda}_k > 0$ against the alternative $H_1 : \hat{\lambda}_k \leqslant 0$. The statistical test can be defined as follows,

$$\hat{t}_k = \frac{\hat{\lambda}_k}{\sqrt{\hat{\varphi}_k / M}} \sim N\big(0, \hat{\varphi}_k\big) \tag{19}$$

Under the null hypothesis H_0: the data-generating process is chaotic, the Lyapunov exponent estimator $\hat{\lambda}_k$ in both cases (global and local context) leads to asymptotically valid inferences in that the associated p-value follows a normal distribution on $N\big(0, \hat{\varphi}_k\big)$. We have estimated the asymptotic variance (φ_k) following the method proposed by Andrews [69]. Thus we have used these results to calculate the standard error of the Lyapunov exponent estimator and investigate its significance. For example, for a significance level of 5%, if we obtain a p-value lower than this level, we would reject the null hypothesis, which would mean that there is no chaotic behaviour.

4. Result and discussion

This section reports the main empirical results of this paper. We have considered four well-known chaotic dynamic systems. The Logistic map, the Gauss map, the Hénon map (in discrete-time) and the Rössler system (in continuous-time), see Table 1. The values of the parameters to obtain those theoretical Lyapunov exponents have been extracted from Devaney [70].

The data given in Table 1 provides the following comments. First, all dynamic systems show a chaotic behaviour for certain parameter values. That is, all of them verify the condition that they have at least one positive Lyapunov exponent. Second, the number of Lyapunov exponents for each dynamic system is equal to the number of dimensions in the state space respectively. For instance, the Gauss map will have only one Lyapunov exponent value since its states space is $\mathbb{R}$ while the Rössler system will have three different Lyapunov exponent values because its states space is $\mathbb{R}^3$.

Third, for k-dimensional deterministic systems with a dimension less than 3 it is only possible to find a chaotic behaviour in discrete-time systems e.g., Logistic map, Gauss map or Hènon map, never in continuous-time systems. We can see this condition in the result provided by the Rössler system: one exponent is negative to guarantee that the system is dissipative, the second exponent is positive verifying the initial-value sensitivity property and the third exponent is zero.

Table 2
Mean square error (MSE) of the largest Lyapunov exponent estimation of one-dimensional chaotic dynamic systems. The MSE values based on the estimation of the largest Lyapunov exponent from direct methods tseriesChaos (D1) and nonlinearTseries (D2) are showed. Those obtained by the Global nerural net (GN) Jacobian indirect methods through the DChaos library are provided. We have developed the new algorithms written in R for local polynomial kernel models (LP) and local neural net kernel models (LN) considering the Norma-2 and QR decomposition procedures.

	Mean Square Error (MSE) of the largest Lyapunov exponent estimation					
Logistic map	$s = 0$	$s = 0.01$	$s = 0.02$	$s = 0.03$	$s = 0.04$	$s = 0.05$
$1 - \textit{Direct} : \textit{tseriesChaos}(D1)$	0.0001220	0.0056643	0.0030120	0.003006	0.0033485	0.0030913
$2 - \textit{Direct} : \textit{nonlinearTseries}(D2)$	0.0802315	0.4765133	0.4814125	0.4815446	0.4790305	0.4830895
$3 - \textit{Indirect} : \textit{N2global}(GN1)$	0.0000324	0.0000382	0.0000691	0.0000994	0.0001314	0.0001532
$4 - \textit{Indirect} : \textit{N2local}(LP1)$	0.0000134	0.0000150	0.0000223	0.0000376	0.0000488	0.0000564
$5 - \textit{Indirect} : \textit{N2local}(LN1)$	0.0000112	0.0000147	0.0000237	0.0000349	0.0000472	0.0000540
$6 - \textit{Indirect} : \textit{QRglobal}(GN2)$	0.0000331	0.0000348	0.0000672	0.0000986	0.0000997	0.0001124
$7 - \textit{Indirect} : \textit{QRlocal}(LP2)$	0.0000152	0.0000197	0.0000264	0.0000359	0.0000422	0.0000516
$8 - \textit{Indirect} : \textit{QRlocal}(LN2)$	0.0000127	0.0000162	0.0000231	0.0000355	0.0000413	0.0000528
Gauss map						
$9 - \textit{Direct} : \textit{tseriesChaos}(D1)$	0.0005270	0.0111180	0.0205349	0.0293853	0.0275621	0.0336681
$10 - \textit{Direct} : \textit{nonlinearTseries}(D2)$	0.0474216	0.1480353	0.1477251	0.1464405	0.1481204	0.1476371
$11 - \textit{Indirect} : \textit{N2global}(GN1)$	0.0000436	0.0000526	0.0000555	0.0000678	0.0000719	0.0000944
$12 - \textit{Indirect} : \textit{N2local}(LP1)$	0.0000170	0.0000199	0.0000247	0.0000388	0.0000460	0.0000507
$13 - \textit{Indirect} : \textit{N2local}(LN1)$	0.0000124	0.0000158	0.0000242	0.0000367	0.0000481	0.0000585
$14 - \textit{Indirect} : \textit{QRglobal}(GN2)$	0.0000618	0.0000656	0.0000672	0.0000782	0.0000817	0.0000924
$15 - \textit{Indirect} : \textit{QRlocal}(LP2)$	0.0000176	0.0000197	0.0000271	0.0000365	0.0000479	0.0000502
$16 - \textit{Indirect} : \textit{QRlocal}(LN2)$	0.0000135	0.0000181	0.0000249	0.0000362	0.0000420	0.0000539

Once the dataset has been described we are interested to know which of the Jacobian indirect methods gives the best results considering also the direct methods. We also want to test if the Jacobian indirect methods provide us (or not) consistent Lyapunov exponent estimators and robustness to the presence of measurement errors. In this sense, we would like to know if as the measurement noise increases, the error committed in obtaining the estimator is amplified with the consequent inaccuracy and inconsistency or instead it is reduced considering some noise-contaminated time series from the chaotic dynamic systems show above.

For this purpose we have compared the theoretical and estimated values of the largest Lyapunov exponent provided by the direct methods, the global Jacobian method (neural net approach) and the local Jacobian method (polynomial kernel approach and neural net kernel model). We have used the R package called *DChaos* (www.CRAN.R-project.org/package=DChaos) to estimate the Global Neural Net Jacobian indirect methods [71]. And we have developed two new algorithms written in R for local polynomial kernel models (LP) and local neural net kernel models (LN) considering the Norma-2 and QR decomposition procedures.

The traditional direct methods used are those proposed by the R packages called *tseriesChaos* [72] and *nonlinearTseries* [73]. These R packages are based on ideas inspired by the time series analysis (TISEAN) project suggested by Hegger et al. [74]. Both implement the direct method provided by Kantz [31]. These R packages are also available at CRAN repository.

We have added to each time series a normal multinomial error term $\varepsilon_t \sim N(0, s)$ with different variance values s. To save CPU time we have set the following parameters: embedding dimension $1 \leqslant m \leqslant 7$, time-delay $\tau = 1$, number of nodes in the single hidden layer $2 \leqslant h \leqslant 10$, order of the polynomial kernel model $1 \leqslant p \leqslant 4$ and sample size $n = 1000$. The classical direct method takes heuristic approaches when estimating the embedding parameters. The global Jacobian methods given by the DChaos package use the statistical approach. We have also follow this same process in our new local algorithms, and we have applied the BIC together with the leave one out cross validation technique.

We have applied the Monte Carlo method. Particularly, we have done 1000 repetitions by different initial conditions when simulating the time series data from the four dynamic systems and six measurement noise levels considered (24000 time series). We have estimated the following models in order to get the results shown inTables 2–3: (i) 63 different global neural nets have been estimated from each 24000 simulated series (1512000 regression models); (ii) 28000 different local polynomial kernel regressions have been estimated from each 24000 simulated series (672000000 regression models); (iii) 63000 different local neural net kernel regressions have been estimated from each 24000 simulated series (1512000000 regression models).

Then the global neural net models, the local polynomial kernel models and the local neural net kernel models has been sorted from lowest to highest BIC values. After that, the best-fitted models with a lower BIC value and their associated embedding parameters has been considered for estimating the 24000 largest Lyapunov exponents. Finally the mean square error (MSE) has been calculated between the theoretical and the estimated value.

We have shown only the results provided by the bootstrap blocking method because as Sandubete and Escot [57] illustrated, on average, gives better results than other blocking methods. The number of bootstrap iterations $B = 1000$. The block length M has been chosen following Shintani and Linton [40] as $M = int\left[c \times (n/\log n)^{1/6}\right]$ with $c = 36.2$ where $int[A]$ signifies the integer part of A. The number of blocks B depends on the sample size n.

Table 3
Mean square error (MSE) of the largest Lyapunov exponent estimation of Multidimensional chaotic dynamic systems. The MSE values based on the estimation of the largest Lyapunov exponent from direct methods tseriesChaos (D1) and nonlinearTseries (D2) are showed. Those obtained by the Global neural Net (GN) Jacobian indirect methods through the DChaos library are also provided. We have developed the new algorithms written in R for local polynomial kernel models (LP) and local neural net kernel models (LN) considering the Norma-2 and QR decomposition procedures.

	Mean Square Error (MSE) of the largest Lyapunov exponent estimation					
Hénon system	$s=0$	$s=0.01$	$s=0.02$	$s=0.03$	$s=0.04$	$s=0.05$
$17-\mathit{Direct}:\mathit{tseriesChaos}(D1)$	0.0005650	0.0067221	0.0092761	0.0100339	0.0141379	0.0189926
$18-\mathit{Direct}:\mathit{nonlinearTseries}(D2)$	0.0121588	0.3133259	0.3145991	0.3115671	0.3226997	0.3178221
$19-\mathit{Indirect}:\mathit{N2global}(GN1)$	0.0000365	0.0000486	0.0000635	0.0000761	0.0000899	0.0000917
$20-\mathit{Indirect}:\mathit{N2local}(LP1)$	0.0000122	0.0000254	0.0000288	0.0000339	0.0000426	0.0000522
$21-\mathit{Indirect}:\mathit{N2local}(LN1)$	0.0000102	0.0000146	0.0000224	0.0000328	0.0000456	0.0000537
$22-\mathit{Indirect}:\mathit{QRglobal}(GN2)$	0.0000318	0.0000451	0.0000589	0.0000601	0.0000866	0.0000932
$23-\mathit{Indirect}:\mathit{QRlocal}(LP2)$	0.0000156	0.0000237	0.0000296	0.0000349	0.0000452	0.0000567
$24-\mathit{Indirect}:\mathit{QRlocal}(LN2)$	0.0000117	0.0000146	0.0000221	0.0000348	0.0000402	0.0000519
Rössler system						
$25-\mathit{Direct}:\mathit{tseriesChaos}(D1)$	0.0004471	0.0049521	0.0063189	0.0072719	0.0127326	0.0174911
$26-\mathit{Direct}:\mathit{nonlinearTseries}(D2)$	0.0398841	0.6412752	0.6388524	0.6396631	0.6451333	0.6499127
$27-\mathit{Indirect}:\mathit{N2global}(GN1)$	0.0002477	0.0003529	0.0005997	0.0006122	0.0009521	0.0019947
$28-\mathit{Indirect}:\mathit{N2local}(LP1)$	0.0001722	0.0001921	0.0002390	0.0002789	0.0003141	0.0003667
$29-\mathit{Indirect}:\mathit{N2local}(LN1)$	0.0001703	0.0001894	0.0002265	0.0002541	0.0003097	0.0003544
$30-\mathit{Indirect}:\mathit{QRglobal}(GN2)$	0.0003168	0.0004891	0.0006070	0.0007155	0.0008190	0.0009268
$31-\mathit{Indirect}:\mathit{QRlocal}(LP2)$	0.0001631	0.0001993	0.0002106	0.0002558	0.0003219	0.0003867
$32-\mathit{Indirect}:\mathit{QRlocal}(LN2)$	0.0001587	0.0001801	0.0002086	0.0002407	0.0003088	0.0003704

QS Kernel with optimal bandwidth has been used for estimating consistently the covariance matrix following Andrews [69]. The MSE values based on the estimation of the largest Lyapunov exponent from the direct methods provided by the tseriesChaos and nonlinearTseries packages are denoted by D1 and D2 respectively. Those obtained by the Global neural net Jacobian indirect methods through the DChaos library are denoted by GN, and the new computational methods are denoted by LP (local polynomial kernel models) and LN (local neural net kernel models) considering the Norma-2 and QR decomposition procedures. The data shown in Tables 2–3 provides the following comments.

First, we can remark that our algorithms are robust to the presence of a (small) measurement noise because the results got are comparable to those which are noise free ($s=0$). Although as the variance of the measurement noise increases, the error committed increases but it is not proportional in any case. Second, the indirect methods provide better estimates than direct methods in all the experiments we have conducted. Third, the algorithms proposed by the tseriesChaos package behave better than those of the nonlinearTseries package. Fourth, local Jacobian indirect methods report better estimates than direct methods and global Jacobian indirect methods when we include a higher noise observing significant differences. Both local nonlinear approaches behave similarly although local neural net kernel models are somewhat better than local polynomial kernel models on average. Then, these results support our working hypothesis, so we have achieved our proposed objectives at the beginning of this paper.

Finally we have focused on testing the reliability of the algorithms used in this paper. For this purpose, we have calculated the *size* and *power* of our hypothesis test. When the null hypothesis is true and we reject it, we make a type I error. The probability of making a type I error is denoted by α, which is the significance level that we set for our hypothesis test. The probability of not rejecting the null hypothesis when it is true (not committing a type I error) is called the size of the test. However, when the null hypothesis is false and we fail to reject it, we make a type II error. The probability of making a type II error is called beta, and is often denoted by β. The probability of rejecting the null hypothesis when it is false (not committing a type II error) is called the power of the test. Remember that our objective is to test the null hypothesis $H_0:\hat{\lambda}_k>0$ against the alternative $H_1:\hat{\lambda}_k\leqslant 0$ considering the statistical test (Eq. (19)).

The rejection percentages displayed provide a measure of the size (Table 4) and the power (Table 5) of the tests at the 5% significance level over 1000 Monte Carlo replications considering the average of all Jacobian indirect algorithms. We have chosen the length $n=50,100,200$ because we want to test the reliability of the algorithms in small sample sizes. Table 4 provides the results from chaotic dynamic systems considering the values of the parameters shown in Table 1. Table 5 gives the results from non-chaotic dynamic systems which have the following parameter set values: $\mu=3.2$ (Logistic map); $\alpha=4.9,\beta=-0.58$ (Gauss map); $a=1.2,b=0.1$ (Hénon system) and $a=0.1,b=0.1,c=7$ (Rössler system).

The results shown in both tables reports the following remarks. First, we can point out that the reliability of the tests is solid at the 5% significance level. The rejection percentages in almost every situation, even for $n=50$, are low when the null hypothesis H_0 is true and high when H_0 is false. Second, the results provide a satisfactory performance in moderate sample sizes. This fact is really important for those researchers who work with short time series data. Third, the empirical size decrease and the empirical power increase as the sample size increase which means that our tests are consistent as well. Fourth, the noise-contaminated data are comparable to those which are noise free. Although as the variance of the

Table 4
Rejection percentages at the 5% significance level from the dynamic systems considered with a chaotic behaviour. These results provide the size of our hypothesis test considering the average of all Jacobian indirect algorithms.

$n = 50$	$s = 0$	$s = 0.01$	$s = 0.02$	$s = 0.03$	$s = 0.04$	$s = 0.05$
Logistic map	2.33	2.54	3.53	3.47	3.23	4.02
Gauss map	2.82	3.89	4.03	4.30	4.63	4.49
Hénon system	2.66	3.50	4.18	4.57	4.64	4.85
Rössler system	3.18	3.52	4.71	4.92	4.52	5.11
$n = 100$						
Logistic map	1.34	1.47	2.62	2.78	3.39	3.70
Gauss map	1.82	2.07	2.36	3.19	3.51	4.23
Hénon system	1.77	1.94	2.41	2.82	3.30	3.65
Rössler system	1.93	2.31	2.69	4.02	3.98	4.57
$n = 200$						
Logistic map	0.45	0.69	1.41	1.56	1.89	2.20
Gauss map	0.58	0.82	0.93	1.21	1.48	2.43
Hénon system	0.41	0.74	1.09	1.72	1.98	2.39
Rössler system	0.73	1.20	1.65	1.97	2.78	3.56

Table 5
Rejection percentages at the 5% significance level from the dynamic systems considered with a non-chaotic behaviour. These results provide the power of our hypothesis test considering the average of all Jacobian indirect algorithms.

$n = 50$	$s = 0$	$s = 0.01$	$s = 0.02$	$s = 0.03$	$s = 0.04$	$s = 0.05$
Logistic map	99.60	99.23	98.57	97.91	96.61	96.12
Gauss map	99.37	98.74	98.59	97.48	97.19	96.80
Hénon system	99.25	99.08	98.70	98.52	97.11	96.92
Rössler system	99.14	98.83	98.30	97.68	97.08	96.71
$n = 100$						
Logistic map	99.78	99.51	98.72	98.19	97.74	97.19
Gauss map	99.49	99.32	98.85	98.69	98.31	97.74
Hénon system	99.54	99.37	98.86	98.66	97.64	97.16
Rössler system	99.30	99.06	98.65	98.21	97.44	97.02
$n = 200$						
Logistic map	99.91	99.72	99.04	98.46	98.13	97.65
Gauss map	99.72	99.57	99.24	99.02	98.78	98.21
Hénon system	99.79	99.50	99.18	98.83	98.18	97.63
Rössler system	99.52	99.17	98.87	98.74	97.95	97.34

measurement noise increases, the errors committed increases but not significantly. Finally, the results shown are fairly robust with respect to the parameter values of the four dynamic systems considered.

5. Conclusion

Methods and techniques related to test the hypothesis of chaos try to quantify the sensitivity to initial conditions estimating the so-called Lyapunov exponents. So quantifying chaos through this kind of quantitative measure is a key point for understanding a chaotic behaviour. This paper describes the main statistical methods for estimating the Lyapunov exponents on a noisy environment by global and local Jacobian indirect algorithms.

The Jacobian indirect approach seeks to estimate the Jacobian of the underlying generating system and then those partial derivatives are used to compute the Lyapunov exponent applying their analytical definition. In this sense, we have illustrated a robust procedure to obtain a consistent estimator of the Lyapunov exponent from the partial derivates by three different nonlinear indirect techniques. First, we have focused on the global neural net approach by approximating the unknown nonlinear system through a feedforward single hidden layer neural network. Those neural net methods provide a well-fit of any unknown linear or nonlinear model. They also have the advantage to allow the analytical derivation, rather than numerically, the Jacobian needed for the estimation of the Lyapunov exponents.

Traditionally the main contributions proposed by the scientific community regarding the local Jacobian indirect methods have focused on local linear approaches. We extend those results by proposing two new nonlinear local approaches based on polynomial kernel regressions and neural net kernel models. In these cases, the estimated partial derivatives will depend on the values taken by the estimated coefficients and the sequence of delayed values evaluated along the whole trajectory for each point. This result is relevant in order to provide better fits to point-wise estimates in this local context because by definition when estimating the Lyapunov exponents through the indirect methods we have to make an evaluation of the Jacobian at each point along the whole trajectory.

We have illustrated the robustness of the algorithms proposed for estimating the Lyapunov exponent from several noise-contaminated time series. We have compared the results provided by the traditional direct methods with those of the Jacobian indirect methods. In this sense we can remark that the indirect methods provide better estimates than direct methods in all the experiments we have conducted. We have only focused on the largest Lyapunov exponent as direct methods do not estimate the full spectrum. We have also shown empirically that our algorithms are robust to the presence of (small) measurement errors because the results obtained are comparable to those which are noise free.

The Jacobian indirect methods seems to perform well for every noisy time series data. The price we have to pay is a greater computational complexity from two points of view, the computing time and the tuning parameters. This fact has a greater effect on both local non-parametric techniques, polynomial kernel regressions and neural net kernel models, than might be expected. Parallelisation is therefore essential for better performance.

Then we have tested the reliability of the algorithms used in this paper. The results have reported a satisfactory performance in moderate sample sizes. This fact is really important for those researchers who work with short time series. The empirical size decreased and the empirical power increased as the sample size increased which means that our tests are consistent as well.

Finally let us remark that the different contributions that compose the Jacobian indirect methods differ in the algorithm used for the estimation of the Jacobian. This fact could motivate the scientific community in this area where new contributions may appear considering new machine learning methods and statistical procedures to obtain robust Lyapunov exponent estimators on a noisy environment in which the presence of just a (small) measurement noise can lead to confusion between a chaotic deterministic system and a purely stochastic one.

Acknowledgments

This work was supported by the Government of Spain (grant RTI2018-094901-B-I00); Complutense University of Madrid (Faculty of Statistical Studies); Camilo José Cela University (Computing and Artificial Intelligence Lab) and the Data Analysis in Social and Gender Studies and Equality Policies Research Group (www.ucm.es/aedipi/).

References

[1] A. Fernández-Díaz, Chaos Theory: Current and Future Research and Applications, McGraw-Hill/Interamericana de España, SL, 2019.
[2] S. Vaidyanathan, C.E. Volos, Advances and Applications in Chaotic Systems, Studies in Computational Intelligence, Vol. 636, Springer, 2016, doi:10.1007/978-3-319-30279-9.
[3] J.C. Sprott, Z. Elhadj, Frontiers in the Study of Chaotic Dynamical Systems with Open Problems, World Scientific Series on Nonlinear Science, Series B, vol. 16, World Scientific, 2011.
[4] K.-S. Chan, H. Tong, Chaos: a statistical perspective, Springer-Verlag, 2001, doi:10.1007/978-1-4757-3464-5.
[5] S. Giannerini, R. Rosa, D.L. Gonzalez, Testing chaotic dynamics in systems with two positive Lyapunov exponents: a bootstrap solution, Int. J. Bifurcation Chaos 17 (01) (2007) 169–182.
[6] K.-S. Chan, Exploration of a Nonlinear World: An Appreciation of Howell Tong's Contributions to Statistics, World Scientific, 2009.
[7] J.Y. Park, Y.-J. Whang, Random walk or chaos: a formal test on the Lyapunov exponent, J. Econom. 169 (1) (2012) 61–74, doi:10.1016/j.jeconom.2012.01.012.
[8] G. Furquim, G. Pessin, B.S. Faiçal, E.M. Mendiondo, J. Ueyama, Improving the accuracy of a flood forecasting model by means of machine learning and chaos theory, Neural Comput. Appl. 27 (5) (2016) 1129–1141.
[9] Z. Lu, B.R. Hunt, E. Ott, Attractor reconstruction by machine learning, Chaos 28 (6) (2018) 061104.
[10] M.A. Sanjuán, Artificial intelligence, chaos, prediction and understanding in science, arXiv preprint arXiv:2003.01771(2020).
[11] M. Faggini, Chaotic time series analysis in economics: balance and perspectives, Chaos 24 (4) (2014) 042101, doi:10.1063/1.4903797.
[12] E. Bradley, H. Kantz, Nonlinear time-series analysis revisited, Chaos 25 (9) (2015) 097610, doi:10.1063/1.4917289.
[13] L. Tang, H. Lv, F. Yang, L. Yu, Complexity testing techniques for time series data: a comprehensive literature review, Chaos Solitons Fractals 81 (2015) 117–135, doi:10.1016/j.chaos.2015.09.002.
[14] V. Oseledec, A multiplicative ergodic theorem. Ljapunov characteristic number for dynamical systems, Trans. Moscow Math. Soc. 19 (1968) 197–231.
[15] A. Lyapunov, The General Problem of the Stability of Motion, Kharkov (1892), (English Transl.: Academic Press, NY, 1966. in Russian
[16] Y.B. Pesin, Characteristic Lyapunov exponents and smooth ergodic theory, Russ. Math. Surv. 32 (4) (1977) 55.
[17] J.P. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors, Rev. Mod. Phys. 57 (1985) 617–656, doi:10.1103/RevModPhys.57.617.
[18] P. Constantin, C. Foiaş, R. Temam, Attractors Representing Turbulent Flows, vol. 314, American Mathematical Soc., 1985.
[19] H.D. Abarbanel, Analysis of Observed Chaotic Data, Springer, 1996.
[20] L.-S. Young, Mathematical theory of Lyapunov exponents, J. Phys. A 46 (25) (2013) 254001.
[21] J.P. Singh, B. Roy, The nature of Lyapunov exponents is (+,+,-,-). is it a hyperchaotic system? Chaos Solitons Fractals 92 (2016) 73–85.
[22] N. Kuznetsov, G. Leonov, T. Mokaev, A. Prasad, M. Shrimali, Finite-time Lyapunov dimension and hidden attractor of the Rabinovich system, Nonlinear Dyn. 92 (2) (2018) 267–285.
[23] N. Kuznetsov, V. Reitmann, Attractor Dimension Estimates for Dynamical Systems: Theory and Computation, Springer, 2020.
[24] K.T. Alligood, T.D. Sauer, J.A. Yorke, Chaos: an Introduction to Dynamical Systems, Springer-Verlag, New York, 1996.
[25] G.A. Leonov, N.V. Kuznetsov, Time-varying linearization and the perron effects, Int. J. Bifurcation Chaos 17 (04) (2007) 1079–1107.
[26] S. Giannerini, R. Rosa, Assessing chaos in time series: statistical aspects and perspectives, Stud. Nonlinear Dyn. Econom. 8 (2) (2004).
[27] N.V. Kuznetsov, T.N. Mokaev, O.A. Kuznetsova, E.V. Kudryashova, The lorenz system: hidden boundary of practical stability and the Lyapunov dimension, Nonlinear Dyn. 102 (2) (2020) 713–732.
[28] R. Gencay, W. Dechert, An algorithm for the n Lyapunov exponents of an n-dimensional unknown dynamical system, Physica D 59 (1) (1992) 142–157, doi:10.1016/0167-2789(92)90210-E.
[29] A. Wolf, J.B. Swift, H.L. Swinney, J.A. Vastano, Determining Lyapunov exponents from a time series, Physica D 16 (3) (1985) 285–317, doi:10.1016/0167-2789(85)90011-9.
[30] M.T. Rosenstein, J.J. Collins, C.J.D. Luca, A practical method for calculating largest Lyapunov exponents from small data sets, Physica D 65 (1) (1993) 117–134, doi:10.1016/0167-2789(93)90009-P.
[31] H. Kantz, A robust method to estimate the maximal Lyapunov exponent of a time series, Phys. Lett. A 185 (1) (1994) 77–87, doi:10.1016/0375-9601(94)90991-1.
[32] H. Kantz, T. Schreiber, Determinism and predictability, Nonlinear Time Ser. Anal. (1997) 42–57.

[33] S. Giannerini, R. Rosa, New resampling method to assess the accuracy of the maximal Lyapunov exponent estimation, Physica D 155 (1–2) (2001) 101–111.
[34] M. Sano, Y. Sawada, Measurement of the Lyapunov spectrum from a chaotic time series, Phys. Rev. Lett. 55 (10) (1985) 1082.
[35] J.P. Eckmann, S.O. Kamphorst, D. Ruelle, S. Ciliberto, Liapunov exponents from time series, Phys Rev A 34 (1986) 4971–4979, doi:10.1103/PhysRevA.34.4971.
[36] R. Brown, P. Bryant, H.D. Abarbanel, Computing the Lyapunov spectrum of a dynamical system from an observed time series, Phys. Rev. A 43 (6) (1991) 2787.
[37] D. Nychka, S. Ellner, A.R. Gallant, D. McCaffrey, Finding chaos in noisy systems, J. R. Stat. Soc. Ser. B Stat. Methodol. 54 (2) (1992) 399–426.
[38] S. Ellner, A. Gallant, D. McCaffrey, D. Nychka, Convergence rates and data requirements for jacobian-based estimates of Lyapunov exponents from data, Phys. Lett. A 153 (6) (1991) 357–363, doi:10.1016/0375-9601(91)90958-B.
[39] M. Shintani, O. Linton, Is there chaos in the world economy? A nonparametric test using consistent standard errors, Int. Econ. Rev. 44 (1) (2003) 331–357.
[40] M. Shintani, O. Linton, Nonparametric neural network estimation of Lyapunov exponents and a direct test for chaos, J. Econom. 120 (1) (2004) 1–33, doi:10.1016/S0304-4076(03)00205-7.
[41] D. Aeyels, Generic observability of differentiable systems, SIAM J. Control Optim. 19 (5) (1981) 595–603.
[42] F. Takens, Detecting Strange Attractors in Turbulence, Springer, Berlin Heidelberg, 1981.
[43] H. Broer, F. Takens, Reconstruction and time series analysis, in: Dynamical Systems and Chaos, Springer, 2011, pp. 205–242.
[44] H. Whitney, Differentiable manifolds, Ann. Math. (1936) 645–680.
[45] H. Tong, K.S. Lim, Threshold autoregression, limit cycles and cyclical data (with discussion), J. R. Stat. Soc. Ser. B 42 (1980) 245–292.
[46] N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, Geometry from a time series, Phys. Rev. Lett. 45 (9) (1980) 712.
[47] R. Mañé, On the dimension of the compact invariant sets of certain non-linear maps, in: Dynamical Systems and Turbulence, Lecture Notes in Mathematics, Springer, 1981, pp. 230–242.
[48] T. Sauer, J.A. Yorke, M. Casdagli, Embedology, J. Stat. Phys. 65 (3–4) (1991) 579–616.
[49] D. Ruelle, F. Takens, On the nature of turbulence, Commun. Math. Phys. 20 (3) (1971) 167–192, doi:10.1007/BF01646553.
[50] H. Kantz, T. Schreiber, Nonlinear Time Series Analysis, vol. 7, Cambridge University Press, 2004.
[51] M.B. Kennel, R. Brown, H.D.I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, Phys. Rev. A 45 (1992) 3403–3411, doi:10.1103/PhysRevA.45.3403.
[52] H. Akaike, Maximum likelihood identification of gaussian autoregressive moving average models, Biometrika 60 (2) (1973) 255–265.
[53] G. Schwarz, Estimating the dimension of a model, Ann. Stat. 6 (2) (1978) 461–464.
[54] G. Claeskens, N.L. Hjort, The focused information criterion, J. Am. Stat. Assoc. 98 (464) (2003) 900–916.
[55] D.F. McCaffrey, S. Ellner, A.R. Gallant, D.W. Nychka, Estimating the Lyapunov exponent of a chaotic system with nonparametric regression, J. Am. Stat. Assoc. 87 (419) (1992) 682–695, doi:10.1080/01621459.1992.10475270.
[56] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Netw. 2 (5) (1989) 359–366, doi:10.1016/0893-6080(89)90020-8.
[57] J.E. Sandubete, L. Escot, Chaotic signals inside some tick-by-tick financial time series, Chaos Solitons Fractals 137 (2020) 109852, doi:10.1016/j.chaos.2020.109852.
[58] J. Fan, I. Gijbels, Local Polynomial Modelling and its Applications: Monographs on Statistics and Applied Probability, vol. 66, CRC Press, 1996.
[59] M.C. Jones, J.S. Marron, S.J. Sheather, A brief survey of bandwidth selection for density estimation, J. Am. Stat. Assoc. 91 (433) (1996) 401–407.
[60] C. Loader, Local Regression and Likelihood, Springer Science & Business Media, 2006.
[61] J. Fan, Q. Yao, Nonlinear Time Series: Nonparametric and Parametric Methods, Springer Science & Business Media, 2008.
[62] N.-B. Heidenreich, A. Schindler, S. Sperlich, Bandwidth selection for kernel density estimation: a review of fully automatic selectors, Adv. Stat. Anal. 97 (4) (2013) 403–433.
[63] Z. Lu, R.L. Smith, Estimating local Lyapunov exponents, Fields Inst. Commun. 11 (1997) 135–151.
[64] Z.-Q. Lu, Estimating Lyapunov exponents in chaotic time series with locally weighted regression, University of North Carolina, 1994 Ph.D. thesis.
[65] Y.-J. Whang, O. Linton, The asymptotic distribution of nonparametric estimates of the Lyapunov exponent for stochastic time series, J. Econom. 91 (1) (1999) 1–42, doi:10.1016/S0304-4076(98)00047-5.
[66] Q. Yao, H. Tong, On prediction and chaos in stochastic systems, Philos. Trans. R. Soc.London Ser. A Phys. Eng. Sci. 348 (1688) (1994) 357–369.
[67] X. Cheng, B. Khomtchouk, N.S. Matloff, P. Mohanty, Polynomial regression as an alternative to neural nets, arXiv preprint arXiv:1806.06850v3 (2019).
[68] R. Gencay, A statistical framework for testing chaotic dynamics via Lyapunov exponents, Physica D 89 (1996) 261–266.
[69] D.W.K. Andrews, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica 59 (3) (1991) 817–858, doi:10.2307/2938229.
[70] R. Devaney, An Introduction to Chaotic Dynamical Systems, Chapman and Hall/CRC Press, 2021.
[71] J.E. Sandubete, L. Escot, DChaos: Chaotic Time Series Analysis, R package version 0.1-5, 2020. https://CRAN.R-project.org/package=DChaos
[72] A.F.D. Narzo, tseriesChaos: Analysis of Nonlinear Time Series, R package version 0.1-13.1, 2019. https://CRAN.R-project.org/package=tseriesChaos
[73] C.A. Garcia, nonlinearTseries: Nonlinear Time Series Analysis, R package version 0.2.6, 2019. https://CRAN.R-project.org/package=nonlinearTseries
[74] R. Hegger, H. Kantz, T. Schreiber, Practical implementation of nonlinear time series methods: the TISEAN package, Chaos 9 (2) (1999) 413–435.