Chaos, Solitons and Fractals 137 (2020) 109852

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

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

Chaotic signals inside some tick-by-tick financial time series

Julio E. Sandubete, Lorenzo Escot*

Faculty of Statistical Studies, Complutense University of Madrid, Av. Puerta de Hierro num 1, 28040 Madrid, Spain

ARTICLE INFO

Article history:
Received 15 November 2019
Revised 19 April 2020
Accepted 27 April 2020
Available online 21 May 2020

Keywords:
Chaos paradox
Tick-by-tick time series
Lagged returns
Non-uniform embedding
Expected lyapunov exponent

ABSTRACT

It has been more than four decades since ideas from chaos began appearing in the literature showing that it is possible to design economic models in regime of chaotic behaviour from a theoretical point of view. However there is no clear evidence that economic time series behave chaotically. So far researchers have found substantial evidence for nonlinearity but relatively weak evidence for chaos. In this paper we propose a possible explanation to this ”chaos model-data paradox”. Our main motivation is that chaos is elusive in financial datasets because of loss of information that occurs when daily quotes are used. This could hinder the detection of chaos in those time series. Chaotic systems are sensitive to initial conditions, so temporal dependence is lost as the chaotic time series are sampled at too long-time intervals, appearing as independent even though they come from a (chaotic) dynamical system. In the case of financial time series, which quotes are continuously traded on markets, the daily sampling may be too long. In order to avoid this problem high-frequency data can be used to detect chaos in financial time series. We have found evidence of chaotic signals inside the 14 tick-by-tick time series considered about some top currency pairs from the Foreign Exchange Market (FOREX). Notice that we do not intend to generalize this finding to all financial series or even to all FOREX series. The main interest of our paper is to illustrate that by choosing a tick-by-tick frequency (instead of a daily one), and with the purpose of preserving the dynamic dependence on the time series, we could find chaos. At least in the 14 specific currency pairs analyzed and during the time intervals considered. Hence we propose take into account all the information available in the financial markets (full sample information on FX rates) instead of daily data. This kind of time series entails several difficulties due to the need to process a huge quantity of information and regarding the reconstruction of the attractor from tick-by-tick time series which are unevenly-spaced. In this sense we have had to implemented new algorithms in order to solve such drawbacks. As far as we know these tick-by-tick financial time series have never been tested for chaos so far.

1. Introduction

It is often said that chaos is an ubiquitous phenomena which is produced everywhere and can be observed from quantum to cosmological scale [1–3]. Particularly, in economics, it has been more than four decades since ideas from chaos began appearing in the literature showing that it is possible to design economic models in regime of chaotic behaviour from a theoretical point of view, see e.g. [4–11]. However there is no clear evidence that economic time series behave chaotically. So far researchers have found substantial evidence for nonlinearity but relatively weak evidence for chaos, see e.g. [12–20]. Hence our interest would be to check why it is usually difficult to detect chaos in economic time series whereas it is apparently easy to show a chaotic behaviour in many theoretical models. This phenomena was called the *chaos model-data paradox* by Brock et al. [21].

In order to explain why chaos is generic in theoretical economics models but elusive in data we will discuss a possible answer to it, at least when studying financial time series. Our main motivation is that chaos is elusive in financial empirical studies because of loss of information that occurs when daily quotes are used. This could hinder the detection of chaos in those time series. Chaotic systems are sensitive to initial conditions, so temporal dependence is lost as the chaotic time series are sampled at too long-time intervals, appearing as independent even though they come from a (chaotic) dynamical system. In the case of financial time series, which quotes are almost continuously changing upwards or backwards on markets, the daily sampling may be too long. In order to avoid this problem high-frequency data can be used to detect chaos in financial time series.

---

* Corresponding author.
E-mail addresses: jsandube@ucm.es (J.E. Sandubete), escot@ucm.es (L. Escot).

Recently some papers have been published looking for nonlinearities and chaos using intra-daily financial time series. For instance, Lo and Lee [22], Aslan and Sensoy [23], Leone and Kwabi [24] focused on testing the hypothesis of efficient markets using high-frequency data; BenSaida [25] analysed the stock performance of the 500 large companies in the United States, S&P 500 index returns, over four different frequencies: weekly, daily, 30-min and 5-min; Anagnostidis and Emmanouilides [26] studied the Athens Exchange Composite Share Price Index (ACSPI) returns over five different frequencies: 10-min, 20-min, 40-min, 50-min and 80-min; Vamvakaris et al. [27] evaluated the S&P 500 index returns as well considering 5-min data. All of them found little evidence to support the presence of chaotic behaviours in their datasets but they found strong evidences of non-linearities. Notice that most of papers published in financial literature so far deal with data equally spaced in time, both in low-frequency as in high-frequency. However financial markets are characterised by its almost continuity on trade, showing a variable time-lapse between each quote. That is, the dataset do not follow a constant rhythm. Therefore this tick-by-tick time series are not separated by a fixed time interval.

In this paper we provide the following contributions: First, instead of using intra-daily data with uniform frequency (1-min, 5-min, 10-min and so on) as the papers mentioned above, we have used intra-daily data with non-uniform frequency (tick-by-tick) which are usually not equally spaced in time. Second, in this sense we have proposed a new algorithm implementing the theorem proposed by Huke and Broomhead [28] who extended the reconstruction theorem when the dynamical system is sampled non-uniformly in time. Their results have allowed us to get the non-uniform delayed-coordinate embedding vectors from the tick-by-tick time series. Third, regarding how to choose the embedding parameters of the reconstruction procedure we have implemented statistical methods based on model selection criteria following Chang and Tong [29,30] instead of the traditional heuristic methods proposed by Abarbanel [18] or Kantz and Schreiber [31].

Fourth, we have estimated the expected Lyapunov exponent using the jacobian indirect method proposed by Eckmann and Ruelle [32] and Gençay and Dechert [33]. Particularly, we have focused on the neural net approach following McCaffrey et al. [34] and Nychka et al. [35] by approximating the unknown non-linear function through a feed-forward single hidden layer network. We have also used different blocking methods proposing a new one based on the bootstrap method. Fifth, these neural net methods are able to fit any unknown linear or non-linear model. This fact has allowed us to be able to calculate analytically rather than numerically the jacobian needed for the estimation of the expected Lyapunov exponents. Sixth, we have proposed new algorithms implementing the results proposed by Shintani and Linton [36] who provided a statistical framework for testing the chaotic hypothesis based on the theoretical asymptotic properties of the neural net estimator.We have developed an R package called *DChaos* which contains the algorithms that we have implemented in this paper. These algorithms are publicly available at the Comprehensive R Archive Network (CRAN) [37].

We have found evidence of chaotic signals inside the 14 tick-by-tick time series considered about some top currency pairs from the Foreign Exchange Market (FOREX). Notice that we do not intend to generalize this finding to all financial series or even to all FOREX series. The main interest of our paper is to illustrate that by choosing a tick-by-tick frequency instead of a daily one, with the purpose of preserving the dynamic dependence on information, we could find chaos. At least in the 14 specific currency pairs analyzed and during the time intervals considered.

The rest of the paper is organised as follow. Section **2** provides a possible explanation to the chaos model-data paradox in finance. Section **3** presents the theoretical framework that we have employed in this paper. Section **4** reports the main results and a discussion of them. Finally, section **5** gives some concluding remarks.

2. Financial data on tick-by-tick time series and chaos

The greatest difficulty in applying chaos theory in empirical studies in economics is related to the quantity and quality of dataset. Unlike experimental sciences where one can carry out experiments and simulations in the laboratory obtaining a large and clean dataset, most economic time series consist of daily, weekly, monthly, quarterly or annual noise contaminated data. The main reason why economic data traditionally were low frequency and discrete was the cost of collection and analysis. Particularly, in financial markets, the emergence of new technologies have had important implications for both the quantity and quality of existing data in finance. Nowadays trades realized by traders are replaced by algorithmic trades executed in an automated way providing a huge amount of data over many intra-daily disaggregated time intervals.

While it is natural to think that with respect to data more is always better, this statement needs to be considered carefully in certain environments. High-frequency financial dataset are still expensive to collect, maintain, and manipulate. These observations are also subject to a wide range of idiosyncratic factors such as non-synchronous trading, intra-day seasonal effects, measurements errors due to bid-ask spreads and even conceptual problems, such as defining a return during an interval in which no trading occur. However, the knowledge that we might have about the market micro-structure depends strongly on access to those high-frequency data and the suitable technological support to store and process that vast amount of data.

Unfortunately, one rarely has the opportunity to access primary databases as most financial markets provide only access to participants, and it is not clear to what extent those data are retained and therefore potentially usable for time series studies. This issue has taken on additional importance with the growth of algorithmic trading. Hence our ability to analyze the dynamic behind financial markets at micro-level is limited mainly by the availability of relevant data. The chance to obtain easily and free high-frequency data on the Foreign Exchange Market market over a sufficiently long time period has made this particular financial market the preferred habitat for studies of nonlinearities and chaos.

The information generated by the interactions between traders who buys (bid orders) and sells (ask orders) those currencies is apparently encoded in frequencies which are not usually equally spaced in time. In our case the frequency of those quotations is tick-by-tick. Each tick will appear when there is a change, upward or downward, in the quotations. That is, in the highest bid or lowest ask orders in the top-of-book. In this paper we have considered 14 tick-by-tick time-series data about some top currency pairs from the Foreign Exchange Market during November 2018. This month has been selected arbitrarily without any economic criteria. The dataset has been obtained from https://truefx.com. This financial market is characterised by being open 24 hours, so it is a market with continuous operations throughout the day.

We have focused on the bid rates top-of-book, tick-by-tick historical market data, with fractional pip spreads in miliseconds detail. In particular we have collected the following currency pairs: the Australian Dollar against Japanese Yen (2,781,632 ticks), the Australian Dollar against US Dollar (2,138,904 ticks), the Canadian Dollar against Japanese Yen (2,115,263 ticks), the Swiss Franc against Japanese Yen (2,579,371 ticks), the Euro against Swiss Franc (2,110,966 ticks), the Euro against British Pound (2,663,370 ticks), the Euro against Japanese Yen (5,909,146 ticks), the Euro against US Dollar (2,963,370 ticks), the British Pound against Japanese Yen (5,023,443 ticks), the British Pound against US Dollar (2,967,520

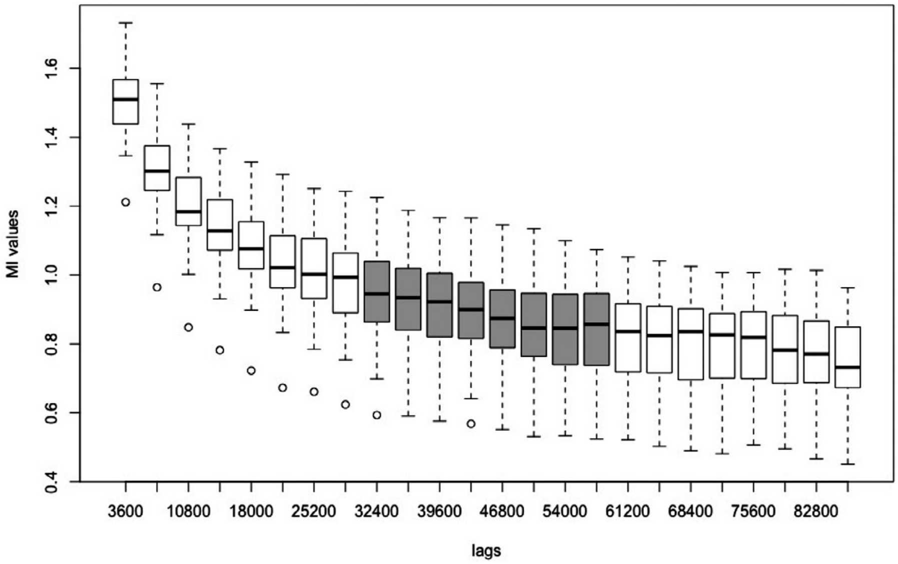

Fig. 1. Boxplot of the Mutual Information values picked in Table 2.

ticks), the New Zealand Dollar against US Dollar (1,398,569 ticks), the US Dollar against Canadian Dollar (1,682,270 ticks), the US Dollar against Swiss Franc (1,497,969 ticks) and the US Dollar against Japanese Yen (2,066,618 ticks).

As we can see the volume of tick-by-tick data is huge when one considers the tick frequency, even taking into account only a month. Thus due to difficult for processing that massive amount of information we have decided to focus on the analysis of FX quotations occurred just during one day. Particularly we have considered two days for the analysis of each FX rate, the highest and the lowest day by volume of tick-by-tick data during November 2018, see Table 1.

As we have pointed out before chaotic systems are sensitive to initial conditions, so temporal dependence in chaotic time series is lost as the series are sampled at too long time intervals, appearing as independent and not chaotic, even though they will in fact come from a chaotic dynamical system. Normally when studying the chaotic behaviour of financial time series, the log returns data is used. In the case of daily time series, the daily returns is usually considered in the literature but if we take into account high-frequency time series we can consider series of higher frequency log returns e.g. 12 hours, 1 h, 10 min, 5 seconds or even tick-by-tick.

Hence to carry out our analysis we need to differentiate the log quotes series in order to get our series of lagged returns by

$$x_t = \log(p_t) - \log\left(p_{t-lag}\right) \tag{1}$$

where $lag \in \mathbb{N}$ is the order of differentiation. If we would be considering the classical daily returns we should use a *lag* equal to 1 day ($lag = 86400$ seconds) but this does not have to be the best option. The optimum lag or order of differentiation should not be too large or too short if we want to preserve the time dependency. On the one hand, a too short lag provides essentially the same information between two consecutive observations and therefore does not allow dynamic dependency to be detected. On the other hand, as a consequence of the initial-value sensitivity property, a too long lag can lead to the loss of information on time dependency. We have followed the heuristic criterion proposed by Fraser and Swinney [38] in order to fix the optimum lag to capture the possible time dependency, chaotic or not, in our tick-by-tick time series. That is, the optimum lag to get our lagged returns will be the first minimum of the mutual information function.

As one can see in both, Table 2 and Fig. 1, we would confirm our hypothesis. That is, we would be losing information if we employ daily frequency returns. We have checked if the first minimum of the mutual information function between $\ln(p_t)$ and $\ln\left(p_{t-lag}\right)$ is lower than 24 hours (86,400 seconds), where p_t is the tick-by-tick bid rates. The mutual information is decreasing as the lag increases. The first minimum of the mutual information function is lower than 86,400 seconds (24 hours) for all currency pairs studied, between $32{,}400 \leq lag \leq 57{,}600$ seconds (9–16 hours), when one takes into account the full sample (tick-by-tick data).

Hence this preliminary analysis establishes that we have to analyse series of financial returns lagged between 9 and 16 hours to optimise the information on time dependency. So, could it be this fact one of the reasons why it is so difficult to find chaos inside our lagged returns?. To answer this question we should detect chaos in the tick-by-tick lagged returns time series in order to justify empirically our hypothesis.

3. How to detect chaotic signals inside tick-by-tick lagged returns time series?

We just showed a possible explanation to the chaos model-data paradox in the field of economics. We have proposed the use of high-frequency data instead of daily data. Indeed, we have used bid rates, tick-by-tick time-series data considering all the information available in the Foreign Exchange Market. Now, we are going to show a procedure to be able to detect chaotic signals inside the tick-by-tick lagged returns time series if there were any. According to the literature, countless techniques have been developed and used to estimate the complexity of time series data, for a review see e.g. Faggini [39], Bradley and Kantz [40], Tang et al. [41]. We have focused on methods derived from chaos theory which estimate the complexity of a dataset through exploring the structure of the attractor. Particularly, we have been interested in the so-called *Lyapunov exponent* (λ_k) as an attractor invariant measure. Quantify chaos through this kind of quantitative measure is a key point for understanding the noisy chaos. Hence our interest will be to test the hypothesis of chaos defined as follows

$$H_0 : \hat{\lambda}_k > 0$$
$$H_1 : \hat{\lambda}_k \leqslant 0$$

Table 1
Samples used in the analysis pertaining to the highest and lowest day in November 2018 by volume of tick-by-tick data are shown.

	AUDJPY	AUDUSD	CADJPY	CHFJPY	EURCHF	EURGBP	EURJPY	EURUSD	GBPJPY	GBPUSD	NZDUSD	USDCAD	USDCHF	USDJPY
Highest day	Nov. 15th	Nov. 15th	Nov. 15th	Nov. 15th	Nov. 15th	Nov. 15th	Nov. 15th	Nov. 15th	Nov. 15th	Nov. 15th	Nov. 01st	Nov. 23th	Nov. 01st	Nov. 15th
Sample size	39,946	69,714	46,900	98,018	69,329	132,588	96,363	88,095	62,188	106,578	25,189	43,694	60,998	50,984
	AUDJPY	AUDUSD	CADJPY	CHFJPY	EURCHF	EURGBP	EURJPY	EURUSD	GBPJPY	GBPUSD	NZDUSD	USDCAD	USDCHF	USDJPY
Lowest day	Nov. 30th	Nov. 19th	Nov. 05th	Nov. 05th	Nov. 05th	Nov. 30th	Nov. 05th	Nov. 19th	Nov. 30th	Nov. 05th	Nov. 05th	Nov. 05th	Nov. 06th	Nov. 30th
Sample size	12,309	40,047	20,635	42,040	41,131	51,610	43,918	53,123	14,923	50,927	14,210	19,774	31,124	21,061

for $k = 1, 2, 3, \ldots$ on a k-dimensional system. Reject the null hypothesis $H_0 : \hat{\lambda}_k > 0$ means lack of chaotic behaviour. That is, the data-generating process does not have a chaotic attractor because of it does not show the property of sensitivity to initial conditions, see Gençay and Dechert and Dechert [33].

In this paper we have considered a noise environment for the following three reasons: (*i*) there are always sources of external random perturbations in the Foreign Exchange Market; (*ii*) there is often a source of noise linked to measurement error in financial datasets; (*iii*) the data-generating process is typically unknown. Hence we have focused on the *expected Lyapunov exponents* from a noisy dynamical system, see Chang and Tong [29,30]. This ergodic measure can be defined as follows. Let $X_t = f(X_{t-1}) + \varepsilon_t$ be a stochastic difference equation where $f : \mathbb{R}^k \to \mathbb{R}^k$, $\{\varepsilon_t\}$ is a sequence of independent and identically distributed d-dimensional random vectors such that ε_t is independent of X_j, $0 \leqslant j \leqslant t-1$, for $t = 1, 2, 3, \ldots, n$. For a k-dimensional system there will be k expected Lyapunov exponents which are given by

$$\begin{aligned}\lambda_k(X_0) &= \lim_{t\to\infty} \frac{1}{t} \mathrm{E}\{\log\left(|Df(X_t)| \cdot \ldots \cdot |Df(X_0)|\right)\} \\ &= \lim_{t\to\infty} \frac{1}{t} \mathrm{E}\left\{\log\left(\left|Df^t(X_0)\right|\right)\right\} \qquad (2)\end{aligned}$$

This relation indicates the expected average rate of divergence or convergence of a random orbit $f^t(X_0)$ starting at the point X_0 where $Df^t(X_0)$ is the jacobian. Since we put ourselves in a experimental context the true dynamics of the system is assumed as unknown. Hence we do not have the advantage of observing directly the state of the system X_t in financial time series let alone knowing the functional form f that generate the dynamic associated with it. Instead of that, there is an observer function $h : \mathbb{R}^k \to \mathbb{R}$ which generates observations as $x_t = h(X_t)$. Therefore it is assumed that all information available is the sequence $\{x_t\}_{t=1}^{n}$ as a time series. Notice that any method for estimating the expected Lyapunov exponent from some observed time series are based previously on the state space reconstruction procedure. Thus we are going to explain how to get the non-uniform delayed-coordinate vectors from our tick-by-tick time series.

3.1. Non-uniform delayed-coordinate embedding

The embedding theorem proposed by Takens [42] provides a framework to reconstruct an unknown dynamical system which gave rise to a given observed scalar time series simply by reconstructing a new state space out of successive values of the time series. The reconstruction theorem assumes that the dynamical system is sampled uniformly in time. There has, however, been increasing interest in situations where observations on the dynamical system are not uniform in time. That is, the data does not come from a series of values measured at uniformly spaced times. For instance, a motivating example would be the case of financial markets. The information generated by the interactions between traders who buys and sells financial instruments such as stocks, bonds, commodities, currencies, derivatives and so on is apparently encoded in frequencies determined by each time a transaction is closed or there is a movement in a bid or ask order which are not equally spaced in time. As we said before this is usually so-called tick-by-tick data. Thus if we observe a particular financial asset i.e., a currency pair on the Foreign Exchange Market market, the measurements of this financial asset are sampling by tick-by-tick intervals which do not follow a constant rhythm.

Can we use this kind of information to build a dynamical system model, and if so, how is it related to the original underlying system?. Huke and Broomhead [28] showed how to extend the reconstruction theorem under certain conditions when the dynamical system is sampled non-uniformly in time. So we can use their

Table 2

Mutual information values between $\ln(p_t)$ and $\ln\left(p_{t-lag}\right)$ of each tick-by-tick time-series data for some top currency pairs. The first minimum is marked in bold type.

lag	AUDJPY	AUDUSD	CADJPY	CHFJPY	EURCHF	EURGBP	EURJPY	EURUSD	GBPJPY	GBPUSD	NZDUSD	USDCAD	USDCHF	USDJPY
3,600	1.42459	1.34625	1.50337	1.50159	1.71316	1.52873	1.21170	1.42573	1.54625	1.45168	1.73233	1.59206	1.53898	1.51053
7,200	1.23655	1.11691	1.29593	1.30134	1.50935	1.35328	0.96424	1.19954	1.37024	1.25494	1.55559	1.38015	1.34643	1.29251
10,800	1.12022	1.00191	1.16691	1.18435	1.40717	1.26871	0.84823	1.08479	1.30135	1.18291	1.43811	1.29774	1.26241	1.17464
14,400	1.05144	0.93061	1.09273	1.11882	1.34925	1.21614	0.78184	1.01184	1.22157	1.12881	1.36683	1.22439	1.17135	1.11558
18,000	0.98691	0.89853	1.05321	1.04852	1.32610	1.13441	0.72233	0.94331	1.17541	1.07617	1.32829	1.19751	1.09473	1.06644
21,600	0.92871	0.83331	1.01367	0.99607	1.29193	1.10482	0.67257	0.89064	1.12633	1.02134	1.26301	1.14892	1.06831	1.01625
25,200	0.90082	0.78429	0.99744	0.96286	1.25098	1.10193	0.66096	0.85904	1.10616	0.97135	1.20728	1.11489	1.02161	1.00179
28,800	0.85602	0.75359	0.96548	0.93156	1.24326	1.07213	0.62382	0.82514	1.05579	0.92491	1.16313	1.07581	0.99364	1.00019
32,400	0.83258	0.69828	0.93776	0.89471	1.22558	**1.04701**	0.59347	0.80185	1.03172	0.91025	1.11786	1.07084	**0.94479**	0.94846
36,000	0.80833	0.65877	0.93379	0.87185	1.18801	1.06716	0.59054	0.78896	0.98381	0.89134	1.08925	1.05505	0.95069	0.91559
39,600	0.77306	**0.64091**	0.93086	**0.86755**	**1.16595**	1.04931	0.57587	**0.75947**	0.97374	**0.86827**	1.03633	1.03991	0.92188	0.89781
43,200	0.73044	0.64211	0.89921	0.86902	1.16637	1.03255	0.56802	0.76269	0.96435	0.87463	0.99281	1.02901	0.91607	0.88747
46,800	0.71662	0.60232	0.87415	0.83121	1.14554	1.02448	0.55146	0.74642	0.94519	0.84249	0.96866	1.02671	0.94222	0.85004
50,400	0.67247	0.57976	0.84612	0.79956	1.13452	1.01389	0.53349	0.72851	0.94274	0.80738	**0.95131**	0.98866	0.92501	**0.83288**
54,000	0.64832	0.55559	**0.83548**	0.77186	1.09961	0.97952	**0.53083**	0.70752	**0.93339**	0.80629	0.95533	**0.95153**	0.90336	0.84541
57,600	**0.62981**	0.52376	0.85609	0.75879	1.07345	0.97612	0.53716	0.71585	0.95414	0.80065	0.94142	0.95926	0.91749	0.85708
61,200	0.63268	0.52171	0.83295	0.73011	1.05239	0.96801	0.53737	0.70701	0.92948	0.77113	0.90274	0.93278	0.90368	0.83642
64,800	0.60392	0.50274	0.82439	0.72112	1.04108	0.96718	0.53542	0.71083	0.93002	0.76172	0.88257	0.92187	0.87971	0.81798
68,400	0.59598	0.48944	0.83611	0.71528	1.02521	0.94166	0.53504	0.67611	0.95333	0.75567	0.86402	0.92123	0.87076	0.80361
72,000	0.59702	0.48123	0.83193	0.72357	1.00721	0.92892	0.54619	0.67727	0.95373	0.73781	0.84931	0.88651	0.82664	0.77131
75,600	0.61451	0.50671	0.81867	0.72218	1.00651	0.93324	0.53669	0.67572	0.93955	0.74966	0.83763	0.88818	0.82161	0.78261
79,200	0.59798	0.49557	0.78218	0.70111	1.01662	0.90463	0.53189	0.67004	0.91406	0.75626	0.84391	0.87592	0.80419	0.75957
82,800	0.58447	0.46626	0.77105	0.69074	1.01375	0.89941	0.53071	0.68439	0.89332	0.74407	0.82951	0.85972	0.82128	0.72281
86,400	0.56083	0.45092	0.72931	0.66752	0.96267	0.91170	0.50383	0.67808	0.87884	0.73124	0.81355	0.84002	0.81609	0.72271

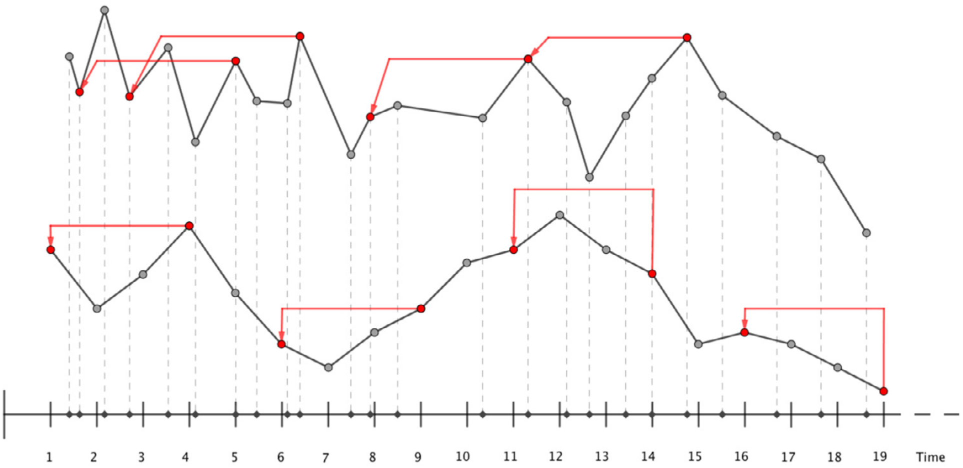

Fig. 2. Uniform (below) and non-uniform (above) time-lapse for delayed-coordinate embedding vectors for any embedding dimension m where the time delay $\tau = 3$.

results to reconstruct the state space from our tick-by-tick time series. We have considered the method of delayed-coordinates proposed by Ruelle and Takens [43] to get the non-uniform delayed-coordinate embedding as follows. Let $\{x_{t_i}\}_{i=1}^{n}$ be our lagged returns time series taking $t_i - t_{i-1} \neq t_s - t_{s-1}$ $\forall i \neq s$. To process this time series we form a sequence of delayed vectors by associating with each time a vector in a reconstructed state space $\mathbb{R}^m$, whose coordinates satisfy the following equation

$$x_{t_i}^m = \left(x_{t_i}, x_{t_i-\tau}, x_{t_i-2\tau}, \ldots, x_{t_i-(m-2)\tau}, x_{t_i-(m-1)\tau}\right) \quad (3)$$

where m is the *embedding dimension* and τ is the *reconstruction time delay* (or lag). As in the uniform case we can construct a vector space whose axes represent all the relevant variables given by

$$x_{t_1}^m = \left(x_{t_1}, x_{t_1-\tau}, \ldots, x_{t_1-(m-2)\tau}, x_{t_1-(m-1)\tau}\right)$$
$$x_{t_2}^m = \left(x_{t_2}, x_{t_2-\tau}, \ldots, x_{t_2-(m-2)\tau}, x_{t_2-(m-1)\tau}\right)$$
$$\vdots$$
$$x_{t_n}^m = \left(x_{t_n}, x_{t_n-\tau}, \ldots, x_{t_n-(m-2)\tau}, x_{t_n-(m-1)\tau}\right)$$

The underlying idea is to make copies of the measured signal with *non-uniform* time-lapse between observations and consider these delayed values as coordinates of a reconstructed state space, retrieved from the data. As in the uniform case the embedding process permits to extract all the relevant information about the unknown underlying dynamical system that generates the time series e.g., the expected Lyapunov exponents defined previously must have approximately the same value in both the true and the reconstructed state space. We have illustrated the phase space reconstruction procedure in both cases for any embedding dimension m where the time delay $\tau = 3$, see Fig. 2.

Notice that a key point to create a suitable reconstruction of the state space is to fix a criteria in order to estimate the embedding parameters (τ and m). Researchers in this area usually estimate them using two different alternatives: a heuristic approach that mostly rely on physical or geometrical arguments and by a statistical approach. Under the heuristic approach regarding the estimation of the time delay τ although there are other criteria, see e.g. Abarbanel [18], Kantz and Schreiber [31], $\tau = 1$ is commonly used following the prescription proposed by Takens and Ruelle [42]. Concerning the embedding dimension m most of the papers published consider the false nearest neighbours criteria proposed by Kennel, Brown and Abarbanel [44]. Another criteria widely used by the scientific community is to estimate the correlation dimension as a proxy of the embedding dimension using the algorithm proposed by Grassberger and Procaccia [45].

The main drawbacks of these heuristic approaches are the following: (*i*) they are not intrinsically statistical; (*ii*) they lead to estimators whose properties are unknown or largely unexplored; (*iii*) they do not take into account the results of any model fit. The alternative proposed by the statistical approach solves those 3 disadvantages. The statistical approach to state space reconstruction can be viewed as a best subset selection problem within the nonparametric regression context as argued Chan and Tong [30]. The idea behind it is to select the embedding parameters, τ and m, that provide the best fit in the estimation of the expected Lyapunov exponents taking into account some information criteria. For instance, the Akaike's information criterion (AIC), the bayesian information criterion (BIC), the Hannan-Quinn information criterion (HQC) or the focused information criterion (FIC). As far as this paper is concerned we will focus in greater detail on model selection procedures instead of heuristic techniques. In any case, we think that the information derived from the heuristic approaches might be still useful and should not be disregarded as a complementary information.

3.2. The expected lyapunov exponent: Estimation & inference

We do not know the true dynamics of the lagged returns time series underlying generating system. For this reason, we will not take into consideration estimation methods that presuppose the knowledge of the equations of the dynamical system. There are basically two kind of methods for estimating the expected Lyapunov exponents from time series. The first one, the so-called *direct* approach which directly measures the growth rate of the divergence between two neighbouring trajectories with an infinitesimal difference in their initial conditions. The direct method was first proposed by Wolf et al. [46], and then revisited by Rosenstein, Collins and De Luca [47], and by Kantz [31,48]. The underlying algorithm is explained in details in Kantz and Schreiber [49] and it is included in the time series analysis (TISEAN) project by Hegger, Kantz and Schreiber [50].

The second one, the so-called *indirect* approach (or Jacobian-based method) which first fit a model to the data based on approximations of the trajectories in the reconstructed state space and then the Jacobian matrices of the model equations are used to compute the expected Lyapunov exponent. The indirect method was first proposed by Eckmann and Ruelle [32], which is based on nonparametric regression methods. Further contributions focuses on two different approaches. Firstly, those who use some kind of local lineal regression, see Sano and Sawada [51], Eckmann et al. [52], Brown, Bryant and Abarbanel [53] or their extension in the form of local polynomial regression proposed by Lu and Smith [54]. Second, other approaches use nonlinear models of neural networks, see McCaffrey et al. [34], Nychka et al. [35], Dechert and Gençay [55], Whang and Linton [56], Shintani and Linton [36,57].

As far as this paper is concerned, we have focused into the jacobian indirect methods using the neural nets approach. Keeping in mind that financial time series are usually noise-contaminated signals and characterised by an erratic and persistent volatility in certain periods, this procedure has the following advantages over all other methods: (*i*) its robustness to the presence of (small) noise; (*ii*) their satisfactory performance in detecting existing nonlinearities on time-series data of moderate sample sizes; (*iii*) allows the estimation of the full spectrum of Lyapunov exponents; (*iv*) the asymptotic distribution of the estimator can be derived allowing the building of formal tests.

Now, let us show how we have implemented the neural nets approach. We have assumed following Gençay and Dechert [33] that there exist a function $g:\mathbb{R}^m \to \mathbb{R}^m$ such that $x_t^m = g\left(x_{t-\tau}^m\right) + \epsilon_t$ where $x_{t_i}^m$ are our non-uniform delayed-coordinate embedding vectors. Under the assumption that the embedding is a homeomorphism, the map g is *topologically conjugate* to the unknown dynamic system f in Eq. 2. This implies that certain dynamical properties of f and g are the same. In our case, the expected Lyapunov exponents of f and g should be the same, so we can focus on estimating the exponents from the map g.

The noisy dynamical system g may be expressed as a matrix by,

$$\begin{pmatrix} x_{t_i} \\ x_{t_i-\tau} \\ \vdots \\ x_{t_i-(m-2)\tau} \\ x_{t_i-(m-1)\tau} \end{pmatrix} = g\begin{pmatrix} x_{t_i-\tau} \\ x_{t_i-2\tau} \\ \vdots \\ x_{t_i-(m-1)\tau} \\ x_{t_i-m\tau} \end{pmatrix} + \epsilon_t$$

$$= \begin{pmatrix} \upsilon\left(x_{t_i-\tau}, x_{t_i-2\tau}, \ldots, x_{t_i-(m-2)\tau}, x_{t_i-(m-1)\tau}, x_{t_i-m\tau}\right) \\ x_{t_i-\tau} \\ \vdots \\ x_{t_i-(m-2)\tau} \\ x_{t_i-(m-1)\tau} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad (4)$$

The jacobian corresponding to the noisy dynamical system g will be the following,

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

Notice that the estimation of the expected Lyapunov exponent by the indirect method is reduced to the estimation of the unknown nonlinear function $\upsilon:\mathbb{R}^m \to \mathbb{R}$. The different contributions that compose the indirect method differ in the algorithm used for the estimation of this function υ in the jacobian (Eq. (5)). We have focused on the neural net approach which is a global nonparametric method that try to estimate the underlying dynamic system without imposing the restriction of local linearity. The neural network (or neural net) estimator of the expected Lyapunov exponent was first proposed by McCaffrey et al.[34] and Nychka et al.[35], and then revisited by Gençay and Dechert [33] and by Shintani and Linton [36,57].

Hornik, Stinchcombe and White [58] showed that any standard feedforward networks with as few as one hidden layer using arbitrary squashing functions are capable of approximating any Borel measurable function from one finite dimensional space to another to any desired degree of accuracy, providing sufficiently many hidden units. In this sense, the feedforward networks are a class of universal approximators. Theoretically, neural nets are expected to perform better than other approximation methods, especially with high-dimensional models, since the approximation form is not so sensitive to the increasing dimension. The results proposed by the authors mentioned above have enabled us to obtain a consistent expected Lyapunov exponent estimator and test the chaotic hypothesis based on the theoretical asymptotic properties of the neural net estimator.

Notice that if we consider the m-dimensional reconstruction vector as defined by Eq. (4)

$$x_{t_i} = \upsilon\left(x_{t_i-\tau}, x_{t_i-2\tau}, \ldots, x_{t_i-(m-2)\tau}, x_{t_i-(m-1)\tau}, x_{t_i-m\tau}\right)$$

The neural network estimator can be obtained by approximating the unknown non-linear function υ through a feed-forward single hidden layer network with a single output by

$$\upsilon \approx \hat{\upsilon} = \Phi_0\left[\hat{\alpha}_0 + \sum_{q=1}^{Q} \hat{\omega}_{q0}\Phi_q\left(\hat{\alpha}_q + \sum_{j=1}^{m} \hat{\omega}_{jq} x_{t_i-j\tau}\right)\right] \quad (6)$$

where $\Phi_0 \in I$, $\hat{\alpha}_0$ is the estimated network bias from input, q is the number of neurones in the single hidden layer, $\hat{\omega}_{q0}$ are the estimated layers connection weights from input to hidden layer, Φ_q is the transfer function which in our case is the logistic function, $\hat{\alpha}_q$ is the estimated network bias from hidden layer, m is the embedding dimension and $\hat{\omega}_{jq}$ are the estimated layers connection weights from hidden layer to output. Notice that in our case the main reason for using neural network models is not to look for the best predictive model but to estimate a model that captures the non-linear time dependence well enough and, additionally, allows us to obtain in an analytical way (instead of numerical) the jacobian functional of the unknown underlying generator system (Eq. (5)). The estimation of this jacobian or partial derivatives will later allow us to contrast our hypothesis of chaos using Eq. (2). We have obtained the partial derivates of the jacobian in Eq. (5) applying the chain rule to Eq. (6) as

$$\frac{\partial \hat{\upsilon}}{\partial x_{t_i-j\tau}} = \Phi'_0(z_0)\sum_{q=1}^{Q} \hat{\omega}_{q0}\Phi'_0(z_q)\hat{\omega}_{jq} \quad (7)$$

where

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

and the estimated partial derivates are given by

$$\hat{D}g = \begin{pmatrix} \frac{\partial \hat{\upsilon}}{\partial x_{t_i-\tau}} & \frac{\partial \hat{\upsilon}}{\partial x_{t_i-2\tau}} & \frac{\partial \hat{\upsilon}}{\partial x_{t_i-3\tau}} & \cdots & \frac{\partial \hat{\upsilon}}{\partial x_{t_i-(m-1)\tau}} & \frac{\partial \hat{\upsilon}}{\partial x_{t_i-m\tau}} \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \quad (8)$$

Once we have estimated the partial derivates as defined in Eq. (7) we would get the neural network estimator of the k-th expected Lyapunov exponent given as

$$\hat{\lambda}_k = \lim_{M\to\infty} \frac{1}{M} \mathrm{E}\left\{\log \mu_k\left(\left|\hat{D}g^M\right|\right)\right\} \quad (9)$$

where $k = 1, 2, 3, \ldots, m$ and μ_k is the k-th largest eigenvalue of $\hat{D}g^M$ where $\hat{D}g^M = \hat{D}g\left(x_{t_M}\right) \cdot \hat{D}g\left(x_{t_{M-1}}\right) \cdot \ldots \cdot \hat{D}g\left(x_{t_1}\right)$ and $\hat{D}g()$ is the estimated jacobian as defined in Eq. (9).

Notice that it is necessary to distinguish between the sample size n used for estimating the partial derivatives of the jacobian in Eq. (7) and the block length M as defined in Eq. (9) which is the number of evaluation points (number of products of the jacobian) used for estimating the k-th expected Lyapunov exponent.

Table 3
Blocking methods for figuring out the neural network estimator of the k-th expected Lyapunov exponent.

Blocking method	Sample size	Block length	Block number	Block subset
Full	n			
Non-overlapping	Blocks	M	$B = n/M$	$1, 2, \ldots, M$ $M+1, M+2, \ldots, 2M$ $\vdots$ $(B-1)M+1, (B-1)M+2, \ldots, BM$
Equally spaced	Blocks	M	$B = n/M$	$1, 1+B, 1+2B, \ldots, 1+(M-1)B$ $2, 2+B, 2+2B, \ldots, 2+(M-1)B$ $\vdots$ $B, 2B, 3B, \ldots, BM$
Bootstrap	Blocks	M	$B = 100$	Randomly

Since the number of evaluation points is less than or equal to n, M can be also understood as the sample size of a subsample. We have taken into account in our algorithms both the full sample and three different methods of subsampling by blocks as we show in Table 3. The first column shows the blocking methods considered. The second, third and fourth column gives the sample size, the block length and the block number of each subsampling method, respectively. The fifth column provides the way in which one picks the position of each element inside the block where each block B corresponds to one row. The bootstrap blocking method takes random samples without replacement by each block.

The asymptotic properties of the nonparametric neural network estimator of the expected Lyapunov exponent $\hat{\lambda}_k$ was derived by Shintani and Linton [36]. They provided a statistical framework for testing the chaotic hypothesis based on the expected Lyapunov exponent and a consistent variance estimator. Their results showed the asymptotic normality of the expected Lyapunov exponent estimators by

$$\sqrt{M}\left(\hat{\lambda}_k - \lambda_k\right) \sim N(0, \varphi_k) \tag{10}$$

where φ_k is the variance of the kth expected Lyapunov exponent. They proved that $\hat{\varphi}_k$ is a consistent variance estimator of φ_k. It can be defined as follows.

$$\hat{\varphi}_k \equiv Var\left(\hat{\lambda}_k\right) = \lim_{M\to\infty} Var\left(\frac{1}{\sqrt{M}}\sum_{t=1}^{M}\eta_{k,t}\right) \tag{11}$$

where M is the subsampling size, that is equal to n for the full sample. The quadratic spectral kernel function $\eta_{k,t}$ is given by $\eta_{k,t} = \xi_{k,t} - \hat{\lambda}_k$ following Shintani and Linton [36]. The parameter $\xi_{k,t}$ is obtained by

$$\xi_{k,t} = \frac{1}{2}\log\mu_k\left(\left|\hat{D}g^t\right|\right) - \frac{1}{2}\log\mu_k\left(\left|\hat{D}g^{t-1}\right|\right) \tag{12}$$

where $t = 1, 2, \ldots, M$. Then a feasible test statistics were introduced and a one-sided test was proposed for the purpose of testing the chaotic hypothesis based on the theoretical asymptotic properties of the neural net estimator. That is, we would know if the expected Lyapunov exponent values are or not statistically significative. Hence our interest will be to test the null hypothesis $H_0 : \hat{\lambda}_k > 0$ against the alternative $H_1 : \hat{\lambda}_k \leqslant 0$. The test statistics can be defined as follows.

$$\hat{t}_k = \frac{\hat{\lambda}_k}{\sqrt{\hat{\varphi}_k/M}} \sim N\left(0, \hat{\varphi}_k\right) \tag{13}$$

Hence we will reject the null hypothesis if $\hat{t}_k \leqslant -z_\alpha$ where z_α is the critical value that satisfies $\Pr[Z \geqslant z_\alpha] = \alpha$ with Z being a standard normal random variable and α is the significance level. Under the null hypothesis H_0 that the data-generating process is chaotic, the neural net estimator $\hat{\lambda}_k$ leads to asymptotically valid inferences in that the associated p-value follows a normal distribution on $N\left(0, \hat{\varphi}_k\right)$. Reject the null hypothesis $H_0 : \hat{\lambda}_k > 0$ means lack of chaotic behaviour. Thus we have used these results to calculate the standard error of the expected Lyapunov exponent estimator and investigate the statistical significance of the sign of the exponents. Now, let us focus on the statistical analysis of the tick-by-tick data considered in this paper.

4. Result and discussion

In this section we provide the main empirical results of this paper. We are going to compare the estimates of the largest expected Lyapunov exponent provided by several blocking methods. We have considered the full sample denoted by $\hat{\lambda}_F$ and three different blocking methods: non-overlapping subsampling $\hat{\lambda}_N$, equally spaced subsampling $\hat{\lambda}_E$ and bootstrap subsampling $\hat{\lambda}_B$, for a review see Table 3. The results have been obtained using different algorithms publicly available at CRAN repository from our R package called DChaos [37]. To save CPU time we have set the embedding dimension $3 \leq m \leq 10$, the time delay $1 \leq \tau \leq 10$, the number of nodes in the single hidden layer $2 \leq Q \leq 10$ and the length of all time-series data is showed in Table 1. In particular 720 different neural nets models have been estimated by each of the 28 tick-by-tick time-series data in order to obtain the results shown in Fig. 3 and Table 4.

Fig. 3 shows the box-plot diagrams of the largest expected Lyapunov exponents derived from all the neural network regressions for each Fx rate based on the bootstrap blocking method. Results pertaining to the highest day by volume of tick-by-tick data are shown above. Those of the lowest day are shown below. As one can see the estimated median value (bold line) of the expected Lyapunov exponent based on 720 different neural nets models estimated are positives in twelve FX rates in both days, see Fig. 3, only the AUDJPY and USDCAD rates among the 14 currency pairs considered have a negative value.

Table 4 shows, for each currency pair, the estimates of the largest expected Lyapunov exponent corresponding to the best-fitted neural net model (with a lower BIC value) for each of the four blocking methods. Let us point out some useful information to understand the results presented in this Table 4. Numbers in parentheses are standard errors. Into brackets [m, τ, Q]. We have indicated in bold those expected Lyapunov exponent with positive values. We have also remarked with an asterisk * those expected Lyapunov exponent that are statistically significant at the 99% confidence level. For the estimation based on blocking methods median values of all used blocks are presented. For the block length M we use $M = int\left[c \times (n/\log n)^{1/6}\right]$ with $c = 36.2$ where $int[A]$ signifies the integer part of A. The number of blocks B depends on the sample size n of each currency pair, see Table 1.

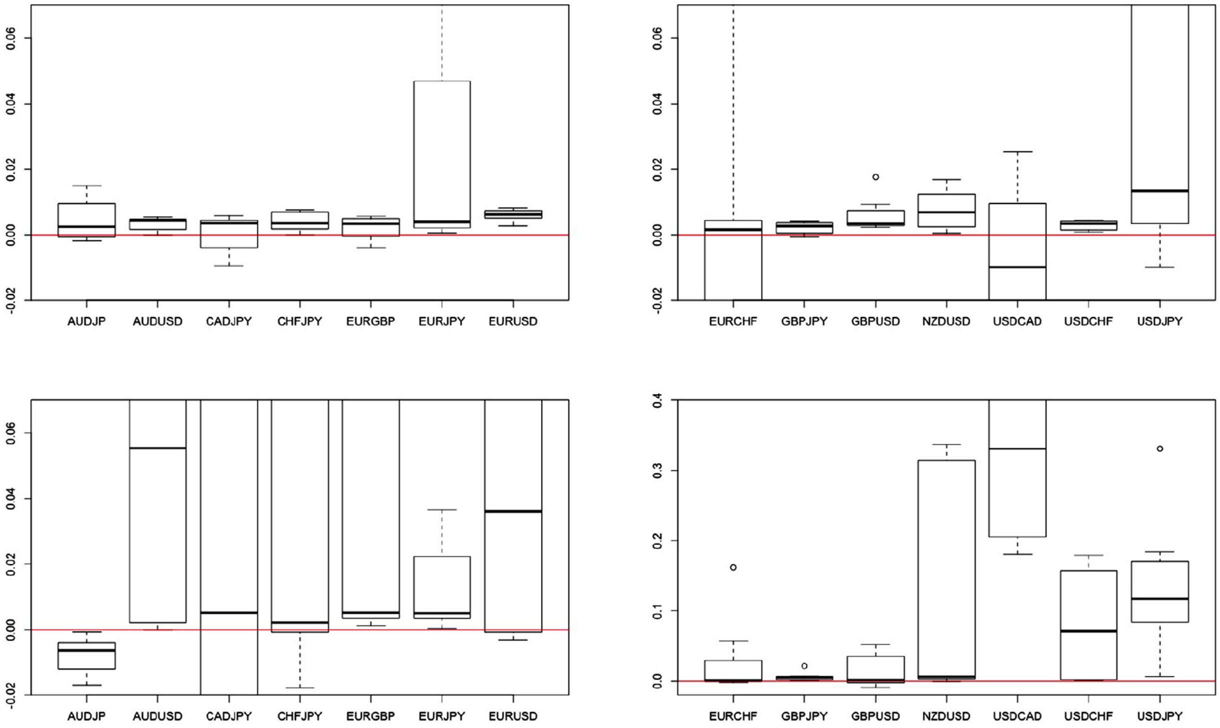

Fig. 3. Box-plot from the expected Lyapunov exponent based on the bootstrap blocking method. We have considered 720 different neural nets models estimated by each currency pair for $3 \leq m \leq 10$, $1 \leq \tau \leq 10$ and $2 \leq Q \leq 10$. Results pertaining to the highest day by volume of tick-by-tick data are shown above. Those of the lowest day are shown below..

Table 4
The expected Lyapunov exponent values based on several blocking methods by each currency pair (F=Full, N=Non-overlapping, E=Equally spaced, B=Bootstrap). Results pertaining to the highest day by volume of tick-by-tick data are shown on the left. Those of the lowest day are shown on the right.

	$[m, \tau, Q]$	$\hat{\lambda}_F$	$\hat{\lambda}_N$	$\hat{\lambda}_E$	$\hat{\lambda}_B$	$[m, \tau, Q]$	$\hat{\lambda}_F$	$\hat{\lambda}_N$	$\hat{\lambda}_E$	$\hat{\lambda}_B$
AUDJPY	[7,1,10]	**0.19177**	**0.18127**	**0.18194**	**0.18199**	[7,1,5]	−0.02013	−0.01061	−0.01057	−0.01043
		(8.1e-04)*	(7.2e-03)*	(8.4e-03)*	(9.0e-03)*		(8.7e-05)	(6.9e-03)	(6.7e-03)	(6.6e-03)
AUDUSD	[8,1,9]	−0.00208	**0.00447**	**0.00445**	**0.00440**	[9,1,8]	**0.16010**	**0.15787**	**0.15488**	**0.15413**
		(1.4e-05)	(4.3e-03)	(4.3e-03)	(4.2e-03)		(1.0e-03)*	(3.2e-03)*	(5.0e-03)*	(4.6e-03)*
CADJPY	[9,1,9]	−0.00094	**0.00572**	**0.00564**	**0.00557**	[10,1,2]	**0.17453**	**0.17184**	**0.17182**	**0.17253**
		(2.0e-05)	(4.2e-03)	(4.1e-03)	(4.0e-03)		(9.0e-04)*	(2.7e-03)*	(6.4e-03)*	(6.1e-03)*
CHFJPY	[8,1,10]	−0.00054	**0.00615**	**0.00673**	**0.00672**	[7,1,10]	−6.9e-06	**0.00529**	**0.00419**	**0.00423**
		(8.3e-04)	(5.4e-03)	(5.3e-03)	(5.2e-03)		(9.4e-05)	(4.5e-03)	(4.3e-03)	(4.6e-03)
EURCHF	[7,1,10]	**0.37030**	**0.36021**	**0.35997**	**0.36013**	[8,1,10]	**0.16506**	**0.16120**	**0.16179**	**0.16177**
		(2.2e-04)*	(7.9e-03)*	(7.7e-03)*	(7.5e-03)*		(3.0e-03)*	(3.2e-03)*	(3.6e-03)*	(3.9e-03)*
EURGBP	[7,1,8]	**0.00172**	**0.00564**	**0.00559**	**0.00554**	[10,1,7]	**0.10923**	**0.10013**	**0.09987**	**0.09998**
		(7.2e-05)*	(2.9e-04)*	(2.9e-04)*	(2.9e-04)*		(9.5e-05)*	(7.1e-03)*	(6.2e-03)*	(6.3e-03)*
EURJPY	[7,1,10]	**0.31488**	**0.30518**	**0.30540**	**0.30588**	[9,1,5]	**0.04076**	**0.03649**	**0.03624**	**0.03687**
		(1.5e-04)*	(5.0e-03)*	(4.9e-03)*	(4.9e-03)*		(2.1e-05)*	(4.0e-03)*	(4.2e-03)*	(4.1e-03)*
EURUSD	[9,1,8]	−0.00080	**0.00759**	**0.00774**	**0.00797**	[9,1,9]	**0.14388**	**0.14191**	**0.14017**	**0.14013**
		(1.4e-05)	(5.0e-03)	(4.8e-03)	(4.8e-03)		(7.5e-04)*	(1.2e-02)*	(1.2e-02)*	(1.2e-02)*
NZDUSD	[7,1,9]	**0.08501**	**0.08241**	**0.08308**	**0.08281**	[8,1,5]	**0.29497**	**0.29438**	**0.29734**	**0.29703**
		(6.4e-04)*	(4.3e-03)*	(7.4e-03)*	(7.9e-03)*		(8.2e-04)*	(3.4e-03)*	(7.6e-03)*	(7.7e-03)*
GBPJPY	[9,1,9]	−0.00236	**0.00372**	**0.00389**	**0.00411**	[10,1,6]	**0.02425**	**0.02116**	**0.02098**	**0.02114**
		(7.5e-05)	(3.7e-03)	(3.6e-03)	(3.6e-03)		(4.8e-05)*	(1.5e-03)*	(1.0e-03)*	(1.1e-03)*
GBPUSD	[10,1,7]	**0.01143**	**0.01664**	**0.01785**	**0.01797**	[9,1,9]	**0.02665**	**0.03286**	**0.02690**	**0.02686**
		(7.5e-04)*	(3.7e-03)*	(3.6e-03)*	(3.6e-03)*		(4.5e-03)*	(5.5e-03)*	(1.2e-03)*	(1.2e-03)*
USDCAD	[9,1,8]	−0.01118	−0.00549	−0.00530	−0.00504	[7,1,10]	**0.61118**	**0.61610**	**0.61349**	**0.61437**
		(8.1e-05)	(4.5e-03)	(4.4e-03)	(4.4e-03)		(2.6e-02)*	(8.6e-03)*	(6.5e-03)*	(6.7e-03)*
USDCHF	[8,1,5]	−0.00058	**0.00427**	**0.00454**	**0.00473**	[10,1,2]	**0.14118**	**0.13634**	**0.13615**	**0.13596**
		(2.5e-05)	(4.8e-03)	(4.7e-03)	(4.7e-03)		(1.1e-03)*	(2.1e-03)*	(5.0e-03)*	(5.1e-03)*
USDJPY	[7,1,10]	**0.36246**	**0.36575**	**0.36713**	**0.36957**	[8,1,10]	**0.3365**	**0.33253**	**0.33091**	**0.33095**
		(5.1e-03)*	(4.8e-03)*	(4.5e-03)*	(4.4e-03)*		(2.6e-04)*	(3.9e-03)*	(4.5e-03)*	(4.4e-03)*

The results shown in Fig. 3 and Table 4 provide the following conclusions. First, if we take into account a sufficiently high embedding dimension m (above 7) the estimated values of the largest expected Lyapunov exponent are positives and significants in both sampled days for the following currency pairs: EURCHF, EURGBP, EURJPY, NZDUSD, GBPUSD, USDJPY. The currency pairs AUDUSD, CADJPY, EURUSD, GBPJPY, USDCHF were positives but only significants on the lowest day by volume of tick-by-tick data, while the CHFJPY rate was positive but not significant in either day.

The AUDJPY and USDCAD rates are the only ones that show negative values on one of the two days. Thus these results would support the hypothesis of chaotic behavior in almost all the analyzed time series. Second, we have found in all cases that the optimum reconstruction delay τ is equal to 1 as Takens and Ruelle suggested [42]. Third, the results provided by the three blocking methods improve the case of the full sample being very similar between them. Although, on average, the bootstrap method gives better results. This result is consistent with the recommendation proposed by Shintani and Linton [36] on the use of blocking methods instead of full sampling when estimating the Lyapunov exponent.

We have been able to detect chaos directly from real-life economic data, at least in the 14 currency pairs analysed and during the days chosen. Hence the results obtained in this paper would support our proposal that chaos is elusive in most financial empirical studies because of loss of information that occurs when daily quotes are used. This could hinder the detection of chaos in those time series. We propose taking into account all the information available in the financial markets (full sample information on quotes) instead of daily data. In this sense we have obtained a possible explanation to this chaos model-data paradox, at least in the 14 tick-by-tick rates and days considered.

5. Conclusion

It has been more than four decades since ideas from chaos began appearing in the literature showing that it is possible to design economic models in regime of chaotic behaviour from a theoretical point of view. However there is no clear evidence that economic time series behave chaotically. So far researchers have found substantial evidence for nonlinearity but relatively weak evidence for chaos. In this paper we have cleared up this chaos model-data paradox in the field of economics showing a possible answer to it, at least when studying financial time series. Chaotic systems are sensitive to initial conditions, so temporal dependence in chaotic time series is lost as the series are sampled at too long time intervals, appearing as independent, even though their underlying generator system could come from a chaotic dynamical system. In the case of financial time series, which quotes are continuously traded on markets, the daily sampling may be too long.

We have used 14 currency pairs on the Foreign Exchange Market to test our suggested solution to the chaos model-data paradox. Our main hypothesis is that chaos is elusive in financial empirical studies because of loss of information that occurs when daily quotes are used. This could hinder the detection of chaos in those time series. We propose taking into account all the information available in the financial markets (full sample information on quotes) instead of daily data. If one wants to detect chaotic signals inside financial time series, we also recommend using lagged returns instead of classical daily returns with the purpose of preserving the dynamic dependence on information. In our case we have calculated them considering a lag or order of differentiation less than a day, between 9 and 16 hours, as indicated by the first minimum of the mutual information function for each currency pair studied.

Most of papers published in financial literature on FOREX so far deal with data equally spaced in time, both in low-frequency (daily) as in high-frequency (1-hour, 30-min, 10-min, 1-min and so on). However this financial market is characterised by its almost continuity, though FX tick-by-tick rates do not follow a constant rhythm. That is, these tick-by-tick time series are not sampling by a fixed time interval. This fact entails several difficulties due to process a huge quantity of information and regarding the attractor reconstruction. In this sense we have implemented an algorithm in order to construct the non-uniform embeddings from our lagged returns time series. This procedure have enabled us to obtain a consistent expected Lyapunov exponent estimator and test the chaotic hypothesis based on the theoretical asymptotic properties of the neural net estimator. These algorithms belong to our R package *DChaos* publicly available at CRAN repository [37].

We conclude this paper asserting that we have been able to detect chaos directly from real financial data on the Foreign Exchange Market. Nevertheless, we have not intended to generalize this finding to all financial series or even to all FOREX series. The main interest of our paper has been to illustrate that by choosing a tick-by-tick frequency instead of a daily one, with the purpose of preserving the dynamic dependence on information, we could find chaos. At least, in the 14 tick-by-tick Fx rates and the days considered. As far as we know these kind of FX tick-by-tick rates have never been tested for chaos. This fact could open up a new line of research in which new contributions may appear considering other financial assets, periods and jacobian indirect methods as convolutional neural nets, local polynomial regression or local neural nets to estimate the expected Lyapunov exponent.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Julio E. Sandubete: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Lorenzo Escot:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgments

This work was supported by the Government of Spain (grant RTI2018-094901-B-I00); the Complutense University of Madrid - Faculty of Statistical Studies (grant CT45/15) and the Data Analysis in Social and Gender Studies and Equality Policies Research Group (www.ucm.es/aedipi/).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi 10.1016/j.chaos.2020.109852.

References

[1] D'Alessio L, Kafri Y, Polkovnikov A, Rigol M. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. Adv Phys 2016;65(3):239–362. doi:10.1080/00018732.2016.1198134.

[2] Dittrich B, Hahn PA, Koslowski TA, Nelson MI. Can chaos be observed in quantum gravity? Phys Lett B 2017;769:554–60. doi:10.1016/j.physletb.2017.02.038.

[3] Hanada M, Shimada H, Tezuka M. Universality in chaos: lyapunov spectrum and random matrix theory. Phys Rev E 2018;97(2):22224. doi:10.1103/PhysRevE.97.022224.

[4] Jensen RV, Urban R. Chaotic price behavior in a non-linear cobweb model. Econ Lett 1984;15(3–4):235–40. doi:10.1016/0165-1765(84)90106-X.

[5] Grandmont J-M. On endogenous competitive business cycles. Econometrica 1985:995–1045. doi:10.2307/1911010.

[6] Chen P. Empirical and theoretical evidence of economic chaos. Syst Dynam Rev 1988;4(1–2):81–108. doi:10.1002/sdr.4260040106.

[7] Kelsey D. The economics of chaos or the chaos of economics. Oxf Econ Pap 1988;40(1):1–31. doi:10.1093/oxfordjournals.oep.a041839.

[8] Hsieh DA. Chaos and nonlinear dynamics: application to financial markets. J Finance 1991;46(5):1839–77. doi:10.1111/j.1540-6261.1991.tb04646.x.

[9] Díaz A. La economía de la complejidad: economía dinámica caótica. McGraw-Hill; 1994. ISBN 9788448118020.

[10] Shone R. Economic dynamics: phase Diagrams and their Economic Application. Cambridge University Press; 2002.

[11] Bruno B, Faggini M, Parziale A, et al. Complexity modelling in economics: the state of the art. Economic Thoughts 2016;5(2):29–43.

[12] LeBaron B. Chaos and nonlinear forecastability in economics and finance. Philos Trans A Math Phys Eng Sci 1994;348(1688):397–404. doi:10.1098/rsta.1994.0099.

[13] Abhyankar A, Copeland LS, Wong W. Nonlinear dynamics in real-Time equity market indices: evidence from the united kingdom. Econ J 1995;105(431):864. doi:10.2307/2235155.

[14] Bask M. Dimensions and Lyapunov exponents from exchange rate series. Chaos Solitons Fractals 1996;7(12):2199–214. doi:10.1016/S0960-0779(96)00080-X.

[15] Barnett WA, Gallant A, Hinich MJ, Jungeilges JA, Kaplan DT, Jensen MJ. A single-blind controlled competition among tests for nonlinearity and chaos. J Econom 1997;82(1):157–92. doi:10.1016/S0304-4076(97)00081-X.

[16] Serletis A, Shintani M. No evidence of chaos but some evidence of dependence in the US stock market? Chaos Solitons Fractals 2003;17(2–3):449–54. doi:10.1016/S0960-0779(02)00387-9.

[17] Park JY, Whang Y-J. Random walk or chaos: a formal test on the lyapunov exponent. J Econom 2012;169(1):61–74. doi:10.1016/j.jeconom.2012.01.012.

[18] Abarbanel HD. Analysis of observed chaotic data. Springer; 1996.

[19] BenSaïda A, Litimi H. High level chaos in the exchange and index markets. Chaos Solitons Fractals 2013;54:90–5. doi:10.1016/j.chaos.2013.06.004.

[20] Lahmiri S. Investigating existence of chaos in short and long term dynamics of moroccan exchange rates. Phys A 2017;465:655–61. doi:10.1016/j.physa.2016.08.024.

[21] Brock W, Hommes C, Schumacher J, Hey C, Hanzon B, Praagman C. Models of complexity in economics and finance. Syst Dyn Econ Financ Models 1997:3–41.

[22] Lo M, Lee C-F. A reexamination of the market efficiency hypothesis: Evidence from an electronic intra-day, inter-dealer fx market. Q Rev Econ Finance 2006;46(4):565–85. doi:10.1016/j.qref.2005.06.002. http://www.sciencedirect.com/science/article/pii/S1062976906000743

[23] Aslan A, Sensoy A. Intraday efficiency-frequency nexus in the cryptocurrency markets. Finance Res Lett 2019. doi:10.1016/j.frl.2019.09.013.

[24] Leone V, Kwabi F. High frequency trading, price discovery and market efficiency in the ftse100. Econ Lett 2019;181:174–7. doi:10.1016/j.econlet.2019.05.022.

[25] BenSaïda A. Noisy chaos in intraday financial data: evidence from the american index. Appl Math Comput 2014;226:258–65. doi:10.1016/j.amc.2013.10.064.

[26] Anagnostidis P, Emmanouilides CJ. Nonlinearity in high-frequency stock returns: evidence from the athens stock exchange. Phys A 2015;421:473–87. doi:10.1016/j.physa.2014.11.056.

[27] Vamvakaris MD, Pantelous AA, Zuev KM. Time series analysis of sp 500 index: a horizontal visibility graph approach. Phys A 2018;497:41–51. doi:10.1016/j.physa.2018.01.010.

[28] Huke JP, Broomhead DS. Embedding theorems for non-uniformly sampled dynamical systems. Nonlinearity 2007;20(9):2205–44. doi:10.1088/0951-7715/20/9/011.

[29] Chan K-S, Tong H. A note on noisy chaos. J R Stat Soc 1994;56(2):301–11. doi:10.1111/j.2517-6161.1994.tb01979.x.

[30] Chan K-S, Tong H. Chaos: a statistical perspective. Springer-Verlag; 2001. doi:10.1007/978-1-4757-3464-5.

[31] Kantz H, Schreiber T. Nonlinear time series analysis, 7. Cambridge university press; 2004.

[32] Eckmann JP, Ruelle D. Ergodic theory of chaos and strange attractors. Rev Mod Phys 1985;57:617–56. doi:10.1103/RevModPhys.57.617.

[33] Gencay R, Dechert W. An algorithm for the n lyapunov exponents of an n-dimensional unknown dynamical system. Phys D 1992;59(1):142–57. doi:10.1016/0167-2789(92)90210-E.

[34] McCaffrey DF, Ellner S, Gallant AR, Nychka DW. Estimating the lyapunov exponent of a chaotic system with nonparametric regression. J Am Stat Assoc 1992;87(419):682–95. doi:10.1080/01621459.1992.10475270.

[35] Nychka D, Ellner S, Gallant AR, McCaffrey D. Finding chaos in noisy systems. J R Stat Soc Series B Stat Methodol 1992;54(2):399–426.

[36] Shintani M, Linton O. Nonparametric neural network estimation of Lyapunov exponents and a direct test for chaos. J Econom 2004;120(1):1–33. doi:10.1016/S0304-4076(03)00205-7.

[37] Sandubete JE, Escot L. Dchaos: chaotic time series analysis. R package version 01–4 2019.

[38] Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. Phys Rev A 1986;33:1134–40. doi:10.1103/PhysRevA.33.1134.

[39] Faggini M. Chaotic time series analysis in economics: balance and perspectives. Chaos 2014;24(4):42101. doi:10.1063/1.4903797.

[40] Bradley E, Kantz H. Nonlinear time-series analysis revisited. Chaos 2015;25(9):097610. doi:10.1063/1.4917289.

[41] Tang L, Lv H, Yang F, Yu L. Complexity testing techniques for time series data: a comprehensive literature review. Chaos Solitons Fractals 2015;81:117–35. doi:10.1016/j.chaos.2015.09.002.

[42] Takens F. Detecting Strange Attractors in Turbulence. Springer Berlin Heidelberg; 1981.

[43] Ruelle D, Takens F. On the nature of turbulence. Commun Math Phys 1971;20(3):167–92. doi:10.1007/BF01646553.

[44] Kennel MB, Brown R, Abarbanel HDI. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys Rev A 1992;45:3403–11. doi:10.1103/PhysRevA.45.3403.

[45] Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. Physica D 1983;9(1–2):189–208.

[46] Wolf A, Swift JB, Swinney HL, Vastano JA. Determining lyapunov exponents from a time series. Phys D 1985;16(3):285–317. doi:10.1016/0167-2789(85)90011-9.

[47] Rosenstein MT, Collins JJ, Luca CJD. A practical method for calculating largest lyapunov exponents from small data sets. Physica D 1993;65(1):117–34. doi:10.1016/0167-2789(93)90009-P.

[48] Kantz H. A robust method to estimate the maximal lyapunov exponent of a time series. Phys Lett A 1994;185(1):77–87. doi:10.1016/0375-9601(94)90991-1.

[49] Kantz H, Schreiber T. Determinism and predictability. Nonlinear Time Ser Anal 1997:42–57.

[50] Hegger R, Kantz H, Schreiber T. Practical implementation of nonlinear time series methods: the tisean package. Chaos 1999;9(2):413–35.

[51] Sano M, Sawada Y. Measurement of the lyapunov spectrum from a chaotic time series. Phys Rev Lett 1985;55(10):1082.

[52] Eckmann JP, Kamphorst SO, Ruelle D, Ciliberto S. Liapunov exponents from time series. Phys Rev A 1986;34:4971–9. doi:10.1103/PhysRevA.34.4971.

[53] Brown R, Bryant P, Abarbanel HD. Computing the lyapunov spectrum of a dynamical system from an observed time series. Phys Rev A 1991;43(6):2787.

[54] Lu Z, Smith RL. Estimating local lyapunov exponents. Fields Inst Commun 1997;11:135–51.

[55] Dechert WD, Gencay R. Lyapunov exponents as a nonparametric diagnostic for stability analysis. J Appl Econ 1992;7(S1):S41–60. doi:10.1002/jae.3950070505.

[56] Whang Y-J, Linton O. The asymptotic distribution of nonparametric estimates of the lyapunov exponent for stochastic time series. J Econom 1999;91(1):1–42. doi:10.1016/S0304-4076(98)00047-5.

[57] Shintani M, Linton O. Is there chaos in the world economy? a nonparametric test using consistent standard errors. Int Econ Rev 2003;44(1):331–57.

[58] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Netw 1989;2(5):359–66. doi:10.1016/0893-6080(89)90020-8.