# Chaotic behavior in financial market volatility

**Article** *in* Journal of Risk · February 2019

DOI: 10.21314/JOR.2018.400

CITATIONS
2

READS
1,195

**4 authors:**

Houda Litimi
University of Jeddah
**14** PUBLICATIONS **99** CITATIONS
SEE PROFILE

Ahmed BenSaïda
Effat University
**26** PUBLICATIONS **243** CITATIONS
SEE PROFILE

Lotfi Belkacem
LaREMFiQ, University of Sousse, Tunisia
**68** PUBLICATIONS **272** CITATIONS
SEE PROFILE

Oussama Abdallah
Université de Rennes 2
**5** PUBLICATIONS **18** CITATIONS
SEE PROFILE

# Chaotic behavior in financial market volatility

Houda Litimi*[1], Ahmed BenSaïda†[1], Lotfi Belkacem‡[1], and Oussama Abdallah§[2]

[1]HEC Sousse - LaREMFiQ Laboratory, University of Sousse, Tunisia
[2]LiRIS Laboratory - EA 7481, University of Rennes, France

**Abstract**

The study of chaotic dynamics in financial time series suffers from the nature of the collected data, which are finite and noisy. Moreover, researchers become less enthusiastic since a large body of literature did not find evidence of chaotic dynamics in financial returns. In this paper, we present a robust method to detect chaos based on the Lyapunov exponent, which is consistent even for noisy and finite scalar time series. To revitalize the debate on nonlinear dynamics in financial markets, we show that the volatility is chaotic. Applications carried on eight daily major volatility indexes support the presence of low-level chaos. Furthermore, the out-of-sample analysis demonstrates the superiority of neural networks, compared to other chaotic maps, in forecasting the market volatility.

*Keywords:* Chaos; Lyapunov exponent; Market risk; Neural network; Nonlinear dynamics
JEL classification: C58, C63, G15, G17

*Email: houda.litimi@gmail.com
†Corresponding author. Email: ahmedbensaida@yahoo.com
‡Email: lotfi.belkacem@yahoo.fr
§Email: oussama.abdallah@univ-rennes2.fr

# 1 Introduction

Studies dealing with financial markets have gained large momentum in the last few decades, and a number of mathematical models have emerged to describe the volatility (*e.g.*, Kim et al., 1998; Eraker et al., 2003; Nakajima and West, 2012; Chan and Grant, 2016). However, these models failed to accurately predict the systems due to considerable temporal variability in the data. Indeed, the stylized facts usually observed in financial returns refrain the researchers from constructing a reliable model that generates good forecasts (Cont, 2001).

After an exuberant body of publications, the study of chaotic dynamics in financial and economic systems has become less appealing to scholars. Despite the diversity of mathematical designs, no empirical evidence of chaos has been supported (Barnett and Hinich, 1992; BenSaïda and Litimi, 2013). Some researchers blame the presence of large noise in financial markets (*e.g.*, Yang et al., 2013; Faggini, 2014); consequently, they hammer away the issue of stochastic modeling (Chan and Grant, 2016). However, the main drawback of existing literature is both conceptual and empirical. In the sense that researchers apply a noise-sensitive tests, that need infinite data, on noisy and finite samples. Furthermore, chaos is investigated solely in the returns, ruling out a very important variable to risk managers and policy makers, which is the volatility.

The literature on deterministic chaotic dynamics in financial processes is essentially focused on market returns (see BenSaïda and Litimi, 2013; Faggini, 2014, for extensive review). Furthermore, there is still no consensus on the adequate way to detect chaos in financial markets. As argued by Faggini (2014), identifying chaos in financial time series is based on several methods, which are, but not limited to: phase space reconstruction (Takens, 1981), Kolmogorov entropy (Grassberger and Procaccia, 1983a), correlation dimension (Grassberger and Procaccia, 1983b), nonlinear forecasting (Farmer and Sidorowich, 1987), false nearest neighbor (Kennel et al., 1992), surrogate data (Theiler et al., 1992), Lyapunov exponent (Kantz, 1994), $0-1$ test for chaos (Gottwald and Melbourne, 2009), multi-step Volterra (Li et al., 2014). However, all these methods suffer from limitations in their fundamental assumptions that are based on infinite sample and noise-free data; and some (correlation dimension and nonlinear forecasting) are designed for more general concept because they include nonlinearity in both deterministic and stochastic systems (BenSaïda and Litimi, 2013). Typically, for financial time series, the problem of small sample size and measurement errors leading to noisy observations have formed the main criticism toward the practical application of these methods (Yang et al., 2013). Moreover, Hu et al. (2005) find that the $0-1$ test for chaos mis-classifies both deterministic and weakly stochastic edge of chaos, and strongly stochastic edge of chaos and weak chaos; thus, it cannot be adequate to financial time series.

Detecting chaotic dynamics in market volatility, as a proxy for risk, proves the existence of an underlying deterministic system, which largely contributes to the existing literature and clarifies the nature and hidden mechanism of financial markets. To the best of our knowledge, this is the first research that investigates the presence of chaos in the volatility of financial markets, and utilizes chaotic maps to forecast the risk, other studies focus on returns.

The objective of our paper is twofold. First, we study the presence of chaotic dynamics in financial market volatility based on the largest Lyapunov exponent (Lyapunov exponent henceforward) with major modifications to account for noise sensitivity and data limited availability. Second, we model the volatility and address the issue of market risk forecasting through chaotic maps. The remainder of this paper is organized as follows. Section 2 presents the employed methodology. Section 3 discusses the numerical procedure. Section 4 tests the presence of chaos in market volatility. Section 5 fits the in-sample volatility through chaotic maps, and exposes the out-of-sample forecasting results. Finally, section 6 concludes the paper.

## 2 Methodology

Chaos was first described by Edward Lorenz at the December 1972 meeting of the American association for the advancement of meteorological and climate science in Washington D.C.. Lorenz advanced the notion of *butterfly-effect* as a metaphorical example to describe nonlinear dynamics in weather evolution and its sensitiveness to initial conditions. A butterfly flapping its wings in India can, in theory, cause a tornado in North America due to nonlinearities in weather processes. The effect of such a small system in producing such a large outcome in a distant region vividly illustrates the impossibility of making predictions in complex systems.

Scientifically speaking, the word *chaos* has a totally different meaning than its general usage as total disaster. With reference to chaos theory, the term chaos indicates an apparent disorder that nevertheless obeys specific rules and is highly sensitive to initial conditions.

Mathematical models can be classified as either deterministic, that does not include randomness, or stochastic (a random process). A deterministic process is when repeated *exactly* in the same way yields *exactly* the same outcome, in contrast to a stochastic process which yields different outcomes when repeated exactly in the same way. One can suggest that deterministic systems exhibit only regular behavior, and if we were able to know the system's current state, we can easily predict its exact future state. However, many deterministic systems exhibit irregular, random-like and unpredictable behavior like the butterfly-effect.

The butterfly-effect is an idiom that connotes the eternal relationship between cause and effect, and that any change in any past event, no matter

how small it is, will result in big changes in the chain-related future events. The practical consequence of the butterfly-effect is that complex systems such as financial markets are difficult to predict over any useful time range. Finite models that attempt to simulate these systems necessarily discard information about the system and the timing of events. These errors are magnified as each unit of time is simulated until the error bound on the result exceeds the limit, such systems are referred to as chaotic. Chaotic systems belong to the class of deterministic dynamical systems that are highly sensitive to initial conditions (Eckmann and Ruelle, 1985). Deterministic chaos is predictable on a short time scale; but on a long time scale, it becomes unpredictable. Chaotic systems behave seemingly irregular, and predictability is lost at a time called *prediction horizon* or *Lyapunov time*, which depends only on the uncertainty of the initial conditions.

The sensitive dependence of chaotic systems on initial conditions implies that any two solution paths with arbitrarily close, yet different initial conditions will diverge exponentially, *i.e.*, for the same chaotic map, small differences in the initial conditions yield large differences in the long-run behavior. Thus, to quantify the main characteristic of a chaotic system, it is possible to compute the rate of divergence known as the *Lyapunov exponent* (Kantz, 1994).[1]

## 2.1 The Lyapunov exponent

Consider two paths in the state space generated respectively by two points $X_0$ and $X_0 + \Delta x_0$ using the same system of equations. The first orbit, generated by $X_0$ is the reference orbit; and the second one is the perturbed orbit that differs from the reference orbit in the initial conditions by $\Delta x_0$. These orbits can be thought as parametric functions of a variable which is related to time $t$. Thus, the separation between the two orbits is also a function of time. Because sensitive dependence can arise only in some portions of a system, this separation is also a function of the location of the initial value and has the form $\Delta x\,(X_0, t)$. In a system with attracting fixed or periodic points, $\Delta x\,(X_0, t)$ decreases asymptotically with time. If a system is unstable, then the orbits diverge exponentially for a while, but eventually settle down. For chaotic points, the function $\Delta x\,(X_0, t)$ behave erratically. The perturbation $\Delta x_0$ created initially between $X_0$ and $X_0 + \Delta x_0$, generates perturbed and unperturbed trajectories, the difference between the two trajectories after $t$ time steps is measured by the Lyapunov exponent $\lambda$ in eq. (1).

[1] Some other techniques to detect chaotic dynamics are also available. For instance, Rubežić et al. (2017) employ a wavelet transform to compute the average wavelet coefficient. Such technique is effective in noisy environments, because wavelets have a denoising effect.

$$\lambda = \lim_{t \to \infty} \frac{1}{t} \ln \frac{|\Delta x\,(X_0, t)|}{|\Delta x_0|} \tag{1}$$

This coefficient measures the average exponential divergence ($\lambda > 0$) or convergence ($\lambda < 0$) rate between nearby trajectories within a time horizon that differ in initial conditions by an infinitesimally small amount. When $\lambda < 0$, the orbit attracts to a stable fixed point, and the system is non-chaotic; and when $\lambda \geqslant 0$, the orbit is unstable and chaotic (see BenSaïda, 2015, for further details).

As argued by BenSaïda (2014), there exist two methods to estimate $\lambda$. The first method proposed by Wolf et al. (1985) based on a direct approach; and the second one by Eckmann and Ruelle (1985) based on the Jacobian approach. The direct approach consists of tracing the exponential divergence of nearby trajectories. In the presence of noise, however, the deterministic divergence is concealed by the noise process on small scales (McCaffrey et al., 1992). This approach is subject to many critics, because it requires infinite series and is sensitive to dynamic noise. The Jacobian approach, on the contrary, can give consistent estimates of the Lyapunov exponent even in presence of noise, as discussed by Nychka et al. (1992); Bailey et al. (1997) and proved by McCaffrey et al. (1992, Theorem 3.1).

Given a scalar time series $\{x_t\}_{t=1}^{T}$, one way to write a noisy chaotic system, using a nonlinear autoregressive model with an additive error, is presented in eq. (2).

$$x_t = f\,(x_{t-L}, x_{t-2L}, \ldots, x_{t-mL}) + \varepsilon_t \tag{2}$$

where $f$ is an unknown smooth function over the real interval that defines the chaotic map, $L$ is the *time delay* that allows skipping samples during the reconstruction, $m$ is the *embedding dimension* or the length of past dependence, and $\varepsilon_t$ is the introduced noise with amplitude $\mathrm{Var}\,(\varepsilon_t) = \sigma_\varepsilon^2$.

Denote $\boldsymbol{X}_t = (x_t, x_{t-L}, \ldots, x_{t-mL+L})' \in \mathbb{R}^m$ and $\boldsymbol{E}_t = (\varepsilon_t, 0, \ldots, 0)' \in \mathbb{R}^m$, then the state-space representation of eq. (2) is defined as:

$$\boldsymbol{X}_t = F\,(\boldsymbol{X}_{t-L}) + \boldsymbol{E}_t$$

such that the map $F : \mathbb{R}^m \mapsto \mathbb{R}^m$, or,

$$F : \begin{bmatrix} x_{t-L} \\ x_{t-2L} \\ \vdots \\ x_{t-mL} \end{bmatrix} \to \begin{bmatrix} x_t = f\,(x_{t-L}, x_{t-2L}, \ldots, x_{t-mL}) + \varepsilon_t \\ x_{t-L} \\ \vdots \\ x_{t-mL+L} \end{bmatrix} \tag{3}$$

Since the dynamics are unknown, BenSaïda (2014) shows that we can reconstruct the map from the observed time series and still preserve the

invariant characteristics of the original unknown map by computing the $m \times m$ Jacobian matrix $\mathbf{J}_t$ of partial derivatives for $F$ evaluated at $\boldsymbol{X}_t$ in eq. (4), *i.e.*, $\mathbf{J}_t = DF(\boldsymbol{X}_t)$.

$$\mathbf{J}_t = \begin{pmatrix} \frac{\partial f}{\partial x_{t-L}} & \frac{\partial f}{\partial x_{t-2L}} & \cdots & \frac{\partial f}{\partial x_{t-(m-1)L}} & \frac{\partial f}{\partial x_{t-mL}} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{4}$$

Therefore, the estimated Lyapunov exponent for a finite sample is derived in eq. (5) (BenSaïda and Litimi, 2013).

$$\hat{\lambda} = \frac{1}{2M} \ln(v_1) \tag{5}$$

where $M$ is the number of equally spaced evaluation points or *block-length*, such that $M \leqslant T$. For accurate estimation of $\lambda$, $M$ should equal $T$. However, smaller values of $M$ increase the estimation time considerably. To offset between accuracy and rapidness with reliable results, BenSaïda and Litimi (2013) propose, without losing efficiency, that $M = \lfloor T^{2/3} \rfloor$, *i.e.*, the integer part of $T^{2/3}$.

$v_1$ is the largest eigenvalue of the matrix $(\mathbf{T}_M \boldsymbol{U}_0)' (\mathbf{T}_M \boldsymbol{U}_0)$, such that $\mathbf{T}_M$ is the product of the Jacobians computed at the evaluation points forming the new subsample of size $M$ (McCaffrey et al., 1992; Nychka et al., 1992; Bailey et al., 1997; Shintani and Linton, 2004). In other words, $\mathbf{T}_M = \prod_{t=0}^{M-1} \mathbf{J}_{M-t}$, and $\boldsymbol{U}_0 = (1, 0, \ldots, 0)'$ is a unit vector to reduce systematic positive bias in the formal estimate of $\lambda$, since $M$ is finite in practice.

It is noteworthy to distinguish between the sample size $T$, which is the total number of observations used to estimate the Jacobian $\mathbf{J}_t$, and the block-length $M$, which corresponds to the sample size of the new subsample or the number of matrices $\mathbf{J}_t$ used to estimate $\lambda$. The estimated $\hat{\lambda}$ converges asymptotically to the largest Lyapunov exponent $\lambda$ as $M \to \infty$, because $\boldsymbol{U}_0$ has zero probability of falling into the subspace corresponding to sub-dominant exponents (Bailey et al., 1997; BenSaïda, 2014).[2]

## 2.2 The chaotic map

To estimate the Lyapunov exponent based on the Jacobian approach, we need to identify the unknown chaotic map $f$. For a scalar time series

[2] Full theoretical development of the Jacobian method and the estimation procedure of the largest Lyapunov exponent $\lambda$ are discussed in McCaffrey et al. (1992); Nychka et al. (1992); Bailey et al. (1997); Shintani and Linton (2004); BenSaïda and Litimi (2013); BenSaïda (2014).

$\{x_t\}_{t=1}^{T}$, BenSaïda and Litimi (2013) propose the neural network to *learn* the process, since it can approximate any smooth, nonlinear function to arbitrary accuracy as the number of hidden layers goes to infinity (Hornik et al., 1989). The basic concept of feed-forward neural net is the propagation of a signal from one layer to another without feedback.

Some researchers (*e.g.*, Garcin and Guégan, 2016; Rubežić et al., 2017, among others) employ a denoising technique based on a wavelet transform to filter the data and estimate the map $f$. Although this approach might be effective for physical systems, where the noise is an intruder to the real pure signal; for financial data, however, microstructure noise is an inherent property to markets. Hence, denoising the data might alter the financial stylized facts, and compromise the true dynamics underlying the time series to be tested. Therefore, we opt for an approach, based on neural networks, to depict the whole process to arbitrary precision.

Let $(x_{t-L}, x_{t-2L}, \ldots, x_{t-mL})$ be the input signal of an $m$-dimensional layer, it is transmitted to a single $q$-dimensional hidden layer with connection weights $\beta_{i,j}$, with $i = 1, \ldots, m$ and $j = 1, \ldots, q$, so the $j^{th}$ cell receives a total signal provided by the foregoing $m$ cells of $\sum_{i=1}^{m} \beta_{i,j} x_{t-iL}$. Moreover, each cell has its own sensitivity called *internal threshold* $\beta_{0,j}$, which is further incorporated as an additional signal. Thus, each cell will receive $\beta_{0,j} + \sum_{i=1}^{m} \beta_{i,j} x_{t-iL}$. The propagation of a signal is governed by an *activation function* of a cell denoted $\Psi$. Theoretically, neural networks support a wide range of activation functions; only a few of them are used by default, the others are available for customization. Bounded activation functions are typically chosen because they have a squashing effect, *i.e.*, they can accept input in any range, and produce output in a strictly limited range. Therefore, the $j^{th}$ cell produces an output signal of the form $\Psi\left(\beta_{0,j} + \sum_{i=1}^{m} \beta_{i,j} x_{t-iL}\right)$. Pal and Mitra (1992) provide a good description of neural networks and multilayer perceptrons. Table 1 reports different activation functions that can be candidates for the neural network estimation. BenSaïda (2014) recommends the hyperbolic tangent function tanh for better performance.

The total $q$-dimensional hidden layer of the neural net, with $\alpha_j$ inter-layers connection weights and $\alpha_0$ network threshold, transmit the entire approximated chaotic map $f$ in eq. (6).

$$x_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \Psi\left(\beta_{0,j} + \sum_{i=1}^{m} \beta_{i,j} x_{t-iL}\right) + \varepsilon_t \tag{6}$$

It is possible to add some other fitting functions to eq. (6), such as polynomial formulation or splines. Nevertheless, the added functions increase the number of coefficients to be estimated, which render any approximation impractical. The estimation procedure consists of minimizing the sum

Table 1: Typical activation functions.

| Function | Form | Range | Explanation |
|---|---|---|---|
| Identity | $x$ | $]-\infty, \infty[$ | The activation level is passed on directly as output. Used in a variety of network types, including linear networks, and the output layer of radial basis function networks. |
| Logistic | $\frac{1}{1+e^{-x}}$ | $[0, \quad 1]$ | An $S$-shaped sigmoid curve. |
| Hyperbolic tangent | $\tanh x$ | $[-1, 1]$ | A sigmoid curve, like the logistic function, except that the output lies in the range $[-1, 1]$. Often performs better than the logistic function because of its symmetry. Ideal for customization of multilayer perceptrons, particularly hidden layers. |
| Sigmoid | $x\frac{1+\frac{\lvert x\rvert}{2}}{2+\lvert x\rvert+\frac{x^2}{2}}$ | $[-1, 1]$ | Another type of sigmoid function. |
| Exponential | $e^{-x}$ | $[0, \infty[$ | Ideal for use with radial units. The combination of radial synaptic function and negative exponential activation function produces units that model a Gaussian function centered at the weight vector. |
| Normalized exponential | $\frac{e^x}{\sum_i e^{x_i}}$ | $[0, \quad 1]$ | Exponential function, with results normalized so that the sum of activations across the layer is 1. Can be used in the output layer of multilayer perceptrons for classification problems, so that the output can be interpreted as probabilities of class membership. |
| Unit sum | $\frac{x}{\sum_i x_i}$ | $[0, \quad 1]$ | Normalizes the output to sum to 1. Used in probabilistic neural networks to allow the output to be interpreted as probabilities. |
| Square root | $\sqrt{x}$ | $[0, \infty[$ | Used to transform the squared distance activation in a self organizing feature map (SOFM) network or cluster network to the actual distance as an output. |
| Sine | $\sin x$ | $[0, \quad 1]$ | Possibly useful when recognizing radially-distributed data. |
| Ramp | $\begin{cases} -1, & \text{if } x \leqslant -1 \\ x, & \text{if } -1 < x < 1 \\ 1, & \text{if } x \geqslant 1 \end{cases}$ | $[-1, 1]$ | A piecewise linear version of the sigmoid function. Has a relatively poor training performance, but fast execution. |
| Step | $\begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geqslant 0 \end{cases}$ | $[-1, 1]$ | Output is either 1 or 0, depending on whether the synaptic value is positive or negative. Can be used to model simple networks such as perceptrons. |

*Note:* This table reports different activation functions that can be candidates for neural network estimation.

of squared noise $S(\theta) = \sum_{t=(mL+1)}^{T} \left[x_t - \hat{f}\left(x_{\underline{t-L}}, \theta\right)\right]^2 = \sum_{t=(mL+1)}^{T} \hat{\varepsilon}_t^2$ to compensate the loss of information engendered from the approximation by the

neural network, where $\theta$ denotes the parameters' vector to be estimated, and $x_{\underline{t-L}}$ is a sequence of past $x_t$, *i.e.*, $x_{\underline{t-L}} = (x_{t-L}, x_{t-2L}, \ldots, x_{t-mL})$. Since the employed activation function $\Psi$ is nonlinear, an obvious estimation algorithm is the nonlinear least squares.[3] To allow pre-sampling, the minimum finite sample size should not be less than $(m\,L + 1)$.

### 2.3 Identifying the orders $(L, m, q)$

The implementation of Takens (1981)'s theorem depends essentially on the embedding dimension $m$ and considers an arbitrary time delay $L$ for infinite and noise-free data. However, for finite and noisy time series, a special care should be attributed to $L$ to keep time dependence between observations. Indeed, if $L$ is too large, then there is a loss of information and subsequent observations may diverge, a condition known as *irrelevance* (Casdagli et al., 1991). Alternatively, if $L$ is too small, the observation vector becomes temporarily close and remote, giving rise to a situation known as *redundance* (Casdagli et al., 1991), where jumps of nearby trajectories do not transmit enough information. It is common to estimate the delay time $L$ from the autocorrelation function (ACF), the average mutual information (AMI), or the correlation integral (CI) (Li et al., 2014). When the data are finite and noisy, the ACF is the preferred approach; still, it shows temporal persistence with gradual decrease in time, and reaches the required value of zero for large $L$ (Li et al., 2014). The main drawback of the CI method is that it needs infinite data (Dhanya and Kumar, 2010). The AMI method selects $L$ based on the first local minimum of the following function:

$$\mathrm{AMI}\,(L) = \sum_{x_t, x_{t-L}} p\,(x_t, x_{t-L}) \ln\left(\frac{p\,(x_t, x_{t-L})}{p\,(x_t)\,p\,(x_{t-L})}\right)$$

where $p\,(x_t, x_{t-L})$ is the joint probability of $x_t$ and $x_{t-L}$, and $p\,(x_t)$ and $p\,(x_{t-L})$ are the individual probabilities of $x_t$ and $x_{t-L}$, respectively. However, the first local minimum may be attained after very large $L$ due to irregularities in the data (Li et al., 2014). As argued by Kim et al. (2009), the AMI method needs large dataset, and may overestimate the time delay.

For the embedding dimension $m$, Takens (1981)'s theorem states that the reconstructed state vector bears the same dynamical properties as the original system for large $m$, two times larger than the unknown dimension of the attractor $d$ $(m > 2d)$. One approach to find the embedding dimension is the *singular system* method of Broomhead and King (1986), where they reconstruct the state space from the lagged observation matrix $\boldsymbol{X}$, by including very large number of lags, even larger than suggested by Takens (1981)'s theorem, and select $m$ as the number of largest eigenvalues of $\mathbf{V} = \boldsymbol{X}'\boldsymbol{X}$,

[3] For linear $\Psi$, such as the identity or step activation functions, it becomes evident to use the linear least squares.

because the largest eigenvectors of $\mathbf{V}$ are the directions in the reconstructed state space showing the variation of the data. However, the number of the largest eigenvectors to retain remains arbitrary. Another approach to find $m$, known as the *false nearest neighbor* (FNN) method, is introduced by Kennel et al. (1992) and improved later by Hegger and Kantz (1999) to avoid spurious results due to noise. Briefly, this method computes the ratio of distances between two neighbor points in the $(m+1)^{th}$ and $m^{th}$ dimensions, if the ratio exceeds a given threshold $u$, the neighbor is false; $m$ is chosen so that the percentage of false neighbors falls to a minimum prefixed value or zero. Hegger and Kantz (1999) suggest a threshold larger than $e^{\lambda L}$ in order for the FNN to reach zero, which remains subjective for a priori unknown $\lambda$.

Theoretically, as the dimension of the hidden layer $q$ goes to infinity, the neural network function can learn any nonlinear function to arbitrary precision (BenSaïda and Litimi, 2013). Nevertheless, too large $q$ increases the computation time exponentially due to large coefficients number; and too low dimension may prevent the neural net from adequately approximating the map.

Nychka et al. (1992) proposed the Bayesian information criterion (BIC) to select the optimal estimated Lyapunov exponent. However, this method penalizes models with higher orders $L$, $m$, and $q$, and influences the test to disregard high-level chaos. Furthermore, as shown by the simulations of BenSaïda and Litimi (2013), the BIC-based approach tends to reject chaos for the same chaotic map when changing the initial conditions.

The choice of the triplet $(L, m, q)$ remains arbitrary in practice. Although, the theory provides some robust methods, they still remain subjective and based on arbitrary criteria. BenSaïda (2014) elaborates a detailed explanation of the strategy to adopt in choosing these parameters without losing efficiency. Indeed, low orders prevent the neural net from efficiently approximating the unknown map $f$, and may overlook high-level chaotic dynamics. On the other hand, high orders add unwanted coefficients structure to the neural net with increasing lagged values of the time series, which results in losing information with the problem of redundant variables, and increases the computational time exponentially.

Therefore, the procedure consists of choosing sufficiently high ceiling triplet $(L, m, q)$, estimate $\lambda$ for each combination of the triplet using nonlinear least square, which yields a total of $L \times m \times q$ Lyapunov exponents, finally, keep the combination of the orders $(L, m, q)$ that maximizes $\hat{\lambda}$. The philosophy behind this method can be explained as follows: if a system is chaotic for any triplet combination that defines the complexity of the neural network, then, negative values of $\hat{\lambda}$ from other order combinations are purely the result of misspecification of the approximating function, which was unable to adequately learn the true dynamics of the process due to inappropriate orders.

## 2.4 Asymptotic distribution of the Lyapunov exponent

The Lyapunov exponent is a constant and fixed measure for a noise-free map (Kantz, 1994). However, when the map is noisy, the accuracy of the estimated Lyapunov exponent is altered and it becomes stochastic (BenSaïda, 2014). Some researches are conducted to infer the asymptotic distribution of the exponent (Lai and Chen, 2003). Following BenSaïda (2014), under the conditions that $F$ is bounded, its Jacobian is also bounded, the noise $\{\varepsilon_t\}_{t=mL+1}^{T}$ is bounded and identically and independently distributed (*i.i.d.*), we have:

$$\sqrt{M}\left(\hat{\lambda}-\lambda\right) \overset{d}{\sim} \mathcal{N}\left(0,\hat{\Sigma}\right), \text{ as } M\to\infty \tag{7}$$

where $\hat{\Sigma}$ is Heteroskedasticity and Autocorrelation Consistent (HAC) estimator of the true variance $\Sigma$ of the *noisy* Lyapunov exponent, and is defined in eq. (8).[4]

$$\hat{\Sigma} = \sum_{j=-M+1}^{M-1} \xi\left(\frac{j}{S_M}\right)\hat{\delta}\left(j\right) \tag{8}$$

where

$$\hat{\delta}\left(j\right) = \frac{1}{M}\sum_{t=|j|+1}^{M} \hat{\eta}_t\hat{\eta}_{t-|j|}$$

such that

$$\hat{\eta}_t = \omega_t - \hat{\lambda}$$

and

$$\begin{cases} \omega_t = & \frac{1}{2}\ln\left[\frac{v_1\left(\mathbf{T}_t'\mathbf{T}_t\right)}{v_1\left(\mathbf{T}_{t-1}'\mathbf{T}_{t-1}\right)}\right], \text{ for } t\geqslant 2 \\ \omega_1 = & \frac{1}{2}\ln\left[v_1\left(\mathbf{T}_1'\mathbf{T}_1\right)\right] \end{cases}$$

$\xi$ and $S_M$ denote a kernel function and a lag-truncation parameter, respectively, employed for HAC estimation since $\eta_t$ are serially correlated and not *i.i.d.* (Andrews, 1991).

**Assumption.** $\xi:\mathbb{R}\mapsto[-1,1]$ is a piecewise continuous function taking the value 1 at zero, symmetric around zero, and has a finite second moment. In other words, the class of kernel is:

$$\Xi = \begin{cases} \xi : \xi\left(0\right)=1, \text{ and } \xi\left(-x\right)=x \quad \forall x\in\mathbb{R}, \text{ with } \int_{-\infty}^{\infty}\xi^2\left(x\right)\mathrm{d}x<\infty; \\ \xi : \text{ is continuous at 0 and all but a finite number of points.} \end{cases}$$

[4] For complete theoretical development, please refer to Shintani and Linton (2004).

The lag-truncation parameter must satisfy $\lim_{M\to\infty} S_M = \infty$ and $\lim_{M\to\infty} \frac{S_M}{M} = 0$. Table 2 reports different kernel functions and their respective optimal lag-truncation parameters. Andrews (1991) has discussed the procedure to estimate $\kappa_1$ and $\kappa_2$ (table 2), which is rather difficult because it depends on the approximating function. BenSaïda and Litimi (2013) have suggested the Quadratic-Spectral with $\kappa_2 = 1$ as an optimal choice among others, since it is valid for any $x \in \mathbb{R}$, and does not disagree in any way with the lag-truncation parameter conditions stated by Andrews (1991).[5]

Table 2: Kernel functions and lag-truncation parameters.

| Kernel | Function | Lag-truncation |
|---|---|---|
| Truncated | $\xi_{TR}(x) = \begin{cases} 1, & \text{for } \lvert x\rvert \leqslant 1 \\ 0, & \text{otherwise.} \end{cases}$ | – |
| Barlett | $\xi_{BT}(x) = \begin{cases} 1-\lvert x\rvert, & \text{for } \lvert x\rvert \leqslant 1 \\ 0, & \text{otherwise.} \end{cases}$ | $S_M = 1.1447\,(\kappa_1 T)^{\frac{1}{3}}$ |
| Parzen | $\xi_{PR}(x) = \begin{cases} 1-6x^2+6\lvert x\rvert^3, & \text{for } 0 \leqslant \lvert x\rvert \leqslant \frac{1}{2} \\ 2(1-\lvert x\rvert)^3, & \text{for } \frac{1}{2} \leqslant \lvert x\rvert \leqslant 1 \\ 0, & \text{otherwise.} \end{cases}$ | $S_M = 2.6614\,(\kappa_2 T)^{\frac{1}{5}}$ |
| Tukey-Hanning | $\xi_{TH}(x) = \begin{cases} \frac{1}{2}(1+\cos \pi x), & \text{for } \lvert x\rvert \leqslant 1 \\ 0, & \text{otherwise.} \end{cases}$ | $S_M = 1.7462\,(\kappa_2 T)^{\frac{1}{5}}$ |
| Quadratic-Spectral | $\xi_{QS}(x) = \frac{25}{12\pi^2 x^2}\left(\frac{\sin\frac{6\pi x}{5}}{\frac{6\pi x}{5}} - \cos\frac{6\pi x}{5}\right)$ | $S_M = 1.3221\,(\kappa_2 T)^{\frac{1}{5}}$ |

*Note:* This table reports different kernel functions and their respective optimal lag-truncation parameters (Andrews, 1991).

## 3 Chaos test and numerical procedure

Chaotic dynamics are detected when the Lyapunov exponent is positive $\lambda \geqslant 0$. In presence of finite and noisy scalar time series, the null hypothesis to test is $H_0 : \lambda \geqslant 0$ against the alternative $H_1 : \lambda < 0$. The rejection of $H_0$ provides strong evidence of the absence of chaotic dynamics. The test statistic to compute is given in eq. (9).

$$\hat{W} = \frac{\hat{\lambda}}{\sqrt{\frac{\hat{\Sigma}}{M}}} \overset{d}{\sim} \mathcal{N}(0,1), \text{ as } M \to \infty \tag{9}$$

The magnitude of $\lambda$ depends on the degree of divergence provided by the map $f$, *i.e.*, as the map gives more divergent outputs, $\lambda$ is expected to increase toward infinity. Similarly, very low chaotic dynamics, with lower

[5] $\lim_{x\to 0} \xi_{QS}(x) = 1$.

divergent outputs, display $\lambda$ almost zero, a condition known as *transition-to-chaos* (BenSaïda, 2015). Conversely, negative values of $\lambda$ indicate the absence of chaotic dynamics in the tested data. Given a significance level $\alpha$, we can accept or reject the null hypothesis of chaotic dynamics. Moreover, small $p$-values cast doubt on the validity of $H_0$. Neural network estimation is based on the *trust-region-reflective* optimization algorithm of Coleman and Li (1996).

The chaos test presented in this paper is a left-tailed test; therefore, the null hypothesis is rejected when $\hat{W} \leqslant -z_\alpha$, where $z_\alpha$ is a critical value that satisfies $\Pr[Z \leqslant -z_\alpha | H_0] = \alpha$, such as $Z$ is a standard normal random variable, and $\alpha$ is the significance level. The $p$-value under the null hypothesis is hence computed in eq. (10).

$$p = \int_{-\infty}^{\hat{W}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \mathrm{d}z \tag{10}$$

BenSaïda and Litimi (2013) have conducted extensive simulations using different stochastic and chaotic models, and concluded that the test have good power in detecting chaos. Furthermore, the test is efficient in providing a clear distinction between stochastic and chaotic dynamics, as opposed to the correlation dimension or BDS test of Brock et al. (1996).

The maximum orders $(L, m, q)$, representing the complexity of the chaotic map, are set to $(5, 6, 5)$, if the null hypothesis of chaos is rejected at this level, we increase the maximum orders to $(10, 12, 10)$ to search for high level chaotic dynamics (BenSaïda and Litimi, 2013). Rejecting the null at higher orders implies the rejection of chaos. Practically, increasing the triplet further is expensive in computer time, since $L \times m \times q$ nonlinear least square estimations are needed to compute the Lyapunov exponent (BenSaïda, 2014). The code can be downloaded from BenSaïda (2015).[6]

# 4 Testing the market volatility

## 4.1 Data and summary statistics

The data are composed of eight daily major volatility indexes, mainly, the VIX (S&P 500 volatility, U.S.), VSTOXX (STOXX 50 volatility, Eurozone), JNIV (Nikkei 225 volatility, Japan), VFTSE (FTSE 100 volatility, U.K.), VDAX (DAX 30 volatility, Germany), VCAC (CAC 40 volatility, France), VAEX (AEX volatility, Netherlands), VSMI (SMI volatility, Switzerland), from January 1, 2001 until December 15, 2016 collected from Datastream. Volatility indexes (divided by 100) are the implied standard deviations of Options on the corresponding market indexes over the next 30 days. Summary statistics are presented in table 3, and volatility indexes are plotted in fig. 1.

[6] Available online at https://github.com/ElsevierSoftwareX/SOFTX-D-15-00007.

Table 3: Summary statistics.

| Volatility index | VIX | VSTOXX | JNIV | VFTSE | VDAX | VCAC | VAEX | VSMI |
|---|---|---|---|---|---|---|---|---|
| Mean | 0.2021 | 0.2496 | 0.2608 | 0.1999 | 0.2438 | 0.2336 | 0.2339 | 0.1938 |
| Minimum | 0.0989 | 0.1160 | 0.1152 | 0.0910 | 0.1165 | 0.0924 | 0.0577 | 0.0924 |
| Maximum | 0.8086 | 0.8751 | 0.9145 | 0.7554 | 0.8323 | 0.7805 | 0.8122 | 0.8490 |
| Std. dev. | 0.0888 | 0.0989 | 0.0911 | 0.0873 | 0.0994 | 0.0905 | 0.1068 | 0.0838 |
| Skewness | 2.1342 | 1.6214 | 2.4789 | 1.8858 | 1.7586 | 1.6733 | 1.7615 | 2.1368 |
| Kurtosis | 9.8453 | 6.2577 | 13.166 | 7.9966 | 6.6000 | 6.7371 | 6.4451 | 9.2923 |
| No. obs. | 4169 | 4169 | 4169 | 4169 | 4169 | 4169 | 4169 | 4169 |

*Note:* This table reports the descriptive statistics of the volatility indexes.

### 4.2 Test results

Chaos test results for market volatility are presented in table 4. For comparison purposes, we further report the results of the $0-1$ test of Gottwald and Melbourne (2009). Chaos is accepted for all studied volatility indexes according to both tests, and the magnitude of the estimated Lyapunov exponents clearly indicates the presence of a low-level chaos in all markets. The volatilities in Eurozone and Netherlands exhibit clear chaotic dynamics, which are detected for low orders $(L, m, q)$. For Switzerland, the Lyapunov exponent is zero, which indicates a *transition-to-chaos* further confirmed by the low orders $(2, 1, 1)$. Nevertheless, for the volatilities in Japan, U.K., and France, chaos is detected only after we increase the complexity of the map to $(10, 12, 10)$.

For the volatilities in U.S. and Germany, although the estimated Lyapunov exponents are negative, their $p$-values are larger then the confidence level $\alpha$ of 5%; therefore, we cannot reject the null hypothesis of chaotic dynamics. Normally, negative $\hat{\lambda}$ indicates that the system is stochastic. However, due to the presence of noise in the data that may conceal the true dynamics, the large uncertainties detected in $\hat{\lambda}$ cast doubt on the absence of chaos, and prevent the test from rejecting the null. Therefore, the volatilities in these markets show low-level chaotic dynamics.

Whereas the $0-1$ test reports merely the presence of chaos, the Lyapunov-based test can also indicate the level and complexity of the chaotic dynamics, *i.e.*, low or high chaos. The estimated Lyapunov exponents for all combinations of the triplet $(L, m, q)$ are plotted in fig. 2.[7] If we detect multiple exact same maximums for each volatility, we keep the orders relative to the parsimonious model, *i.e.*, with lower triplet sum.

[7] Table 4 reports only the maximums of these values for each data.

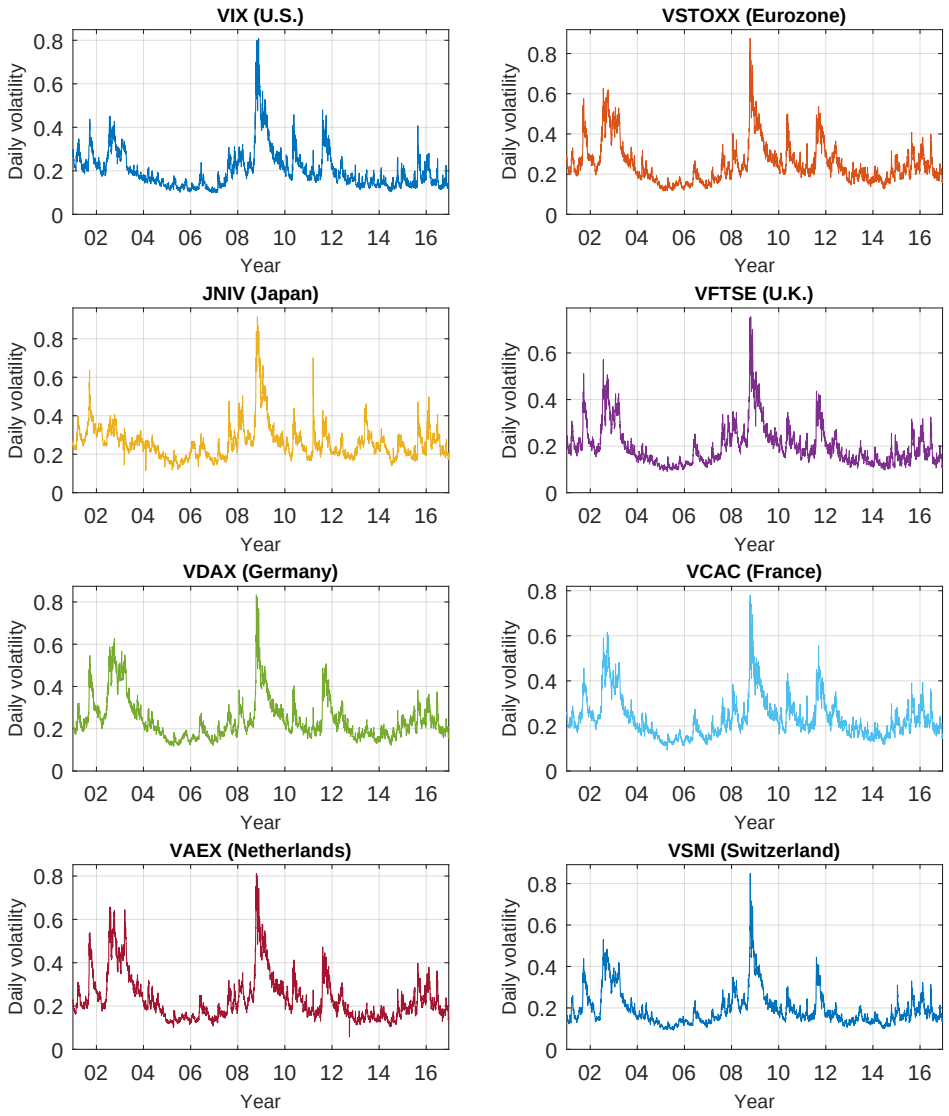


Figure 1: Volatility indexes.

# 5 Modeling and forecasting the chaotic volatility

In this section, (1) we fit some chaotic maps to the in-sample volatility, and (2) we compare the performance of each chaotic model in forecasting the one-step-ahead volatility.

## 5.1 Modeling the chaotic volatility

Chaotic modeling in finance focuses mainly on the behavior of *fundamentalists*, who believe that asset prices are determined solely by the efficient

Table 4: Chaos test results.

| Volatility index | VIX | VSTOXX | JNIV | VFTSE | VDAX | VCAC | VAEX | VSMI |
|---|---|---|---|---|---|---|---|---|
| $\hat{\lambda}$ | -0.0021[‡] [0.1584] | 0.0004[†] [0.5532] | 0.0008[†] [0.6163] | 0.0062[†] [0.7887] | -0.0004[‡] [0.4460] | 0.0029[†] [0.8364] | 0.0670[†] [1.0000] | 0.0000[†] [0.5059] |
| $(L, m, q)$ | $(1, 4, 1)$ | $(1, 4, 1)$ | $(1, 11, 5)$[§] | $(4, 12, 10)$[§] | $(1, 4, 1)$ | $(1, 11, 6)$[§] | $(1, 4, 1)$ | $(2, 1, 1)$ |
| Accepted hypothesis | $H_0$ Chaos | $H_0$ Chaos | $H_0$ Chaos | $H_0$ Chaos | $H_0$ Chaos | $H_0$ Chaos | $H_0$ Chaos | $H_0$ Chaos |
| $0-1$ test[¶] | 0.9898 Chaos | 0.9868 Chaos | 0.9933 Chaos | 0.9880 Chaos | 0.9794 Chaos | 0.9915 Chaos | 0.9717 Chaos | 0.9706 Chaos |

*Note:* This table reports the chaos test results of volatility indexes. $p$-values are in brackets.

[†] The estimated Lyapunov exponent is positive and its $p$-value is greater than the significance level $\alpha$ of 5%; therefore, the null hypothesis of chaotic dynamics cannot be rejected.

[‡] Although the estimated Lyapunov exponent is negative, its $p$-value is greater than the significance level $\alpha$ of 5%; therefore, the null hypothesis of chaotic dynamics cannot be rejected.

[§] Chaos is not detected for maximum orders of $(5, 6, 5)$, yet, it is detected for higher maximum orders of $(10, 12, 10)$, which indicates a higher complexity in the true chaotic map.

[¶] The $0-1$ test reports the asymptotic growth rate $K_c$. When $K_c \to 1$, this indicates the presence of chaos; and when $K_c \to 0$, this indicates regular dynamics (Gottwald and Melbourne, 2009).

market hypothesis; and the *chartists*, who believe that asset prices may be predicted by technical analysis of past prices. Brock and Hommes (1998), for instance, show that routes to chaos may arise in the cobweb model and the capital asset pricing model when introducing the heterogeneous beliefs, with different groups of traders having different expectations about future prices. Guégan (2009) proposes a 2-factor Gegenbauer process to analyze the long-term behavior of a financial variable. Other examples of deterministic economic models, with non-periodic fluctuations, have been developed by Day (1992), among others.

A major problem when studying chaotic dynamics in finance is that the data, which we want to infer the underlying process, have a unique trajectory $(x_1, \ldots, x_T)$. Indeed, a single trajectory is not sufficient to deduce the possible patterns. Unless we have a pre-defined theoretical model, we limit our estimating functions to one-dimensional chaotic maps. Therefore, we select two more candidates beside the neural net: the logistic map in eq. (11) and the Gaussian map (or Gauss iterated map) in eq. (12) (Hilborn, 2000, p. 192). Where $\varepsilon_t$ represents the added noise to the chaotic map, assumed to be normally and independently distributed. Note that Li et al. (2011) claim that feed-forward multilayer neural networks are chaotic, and Bahi et al. (2012) provide a rigorous proof under some conditions.

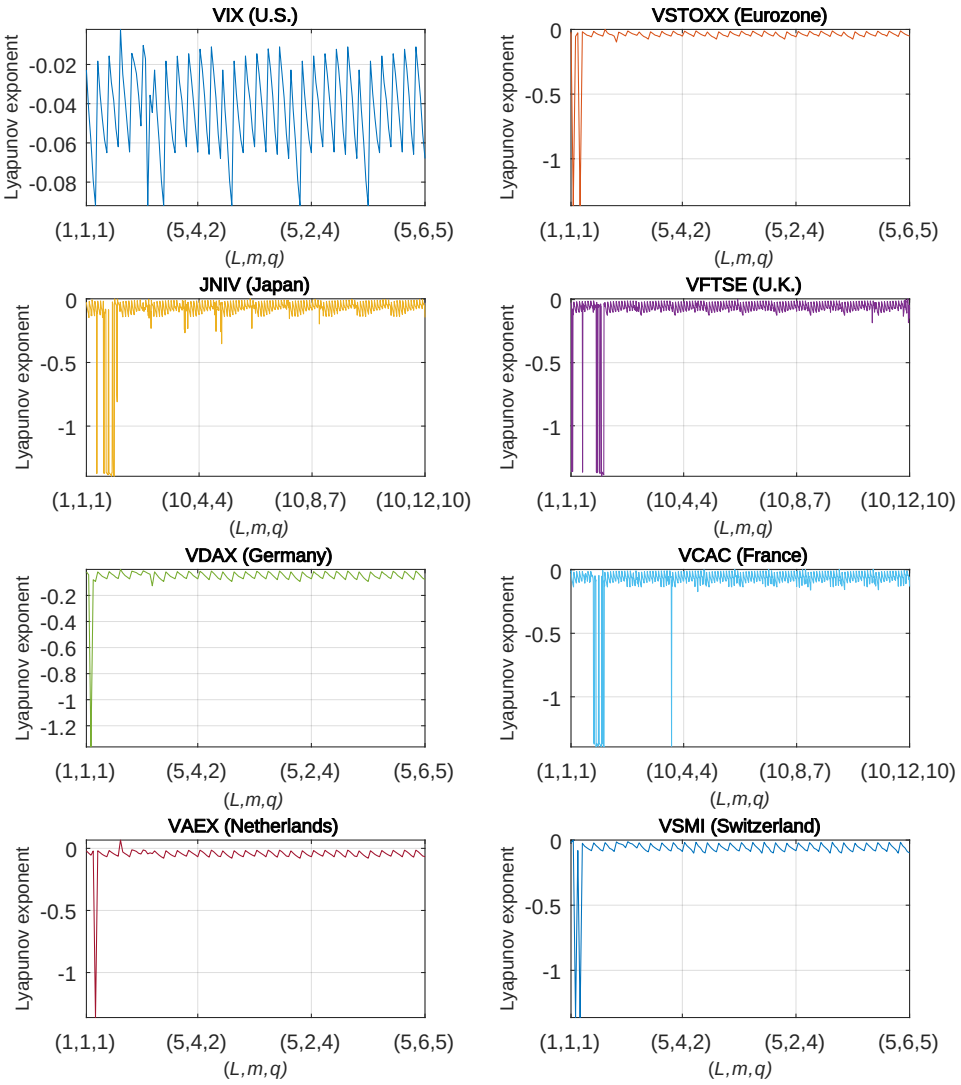


Figure 2: Spectrum of estimated Lyapunov exponents $\hat{\lambda}$.

$$x_{t+1} = \mu\, x_t\,(1 - x_t) + \varepsilon_t, \qquad \text{where } \mu \in [0,4] \text{ and } x_t \in [0,1] \tag{11}$$

$$x_{t+1} = e^{-\alpha x_t^2} + \beta + \varepsilon_t, \qquad \text{where } \alpha, \beta \in \mathbb{R} \text{ and } x_t \in \mathbb{R} \tag{12}$$

The logistic map is defined on the interval $[0,1]$, which makes it inappropriate for general applications. However, since our studied volatility falls inside this domain, we report the estimation results of this map for illustrative purpose. The fitted volatility under each chaotic map is plotted along side the actual values in fig. 3.

Graphically, the logistic map fails to fit periods of high volatility, especially

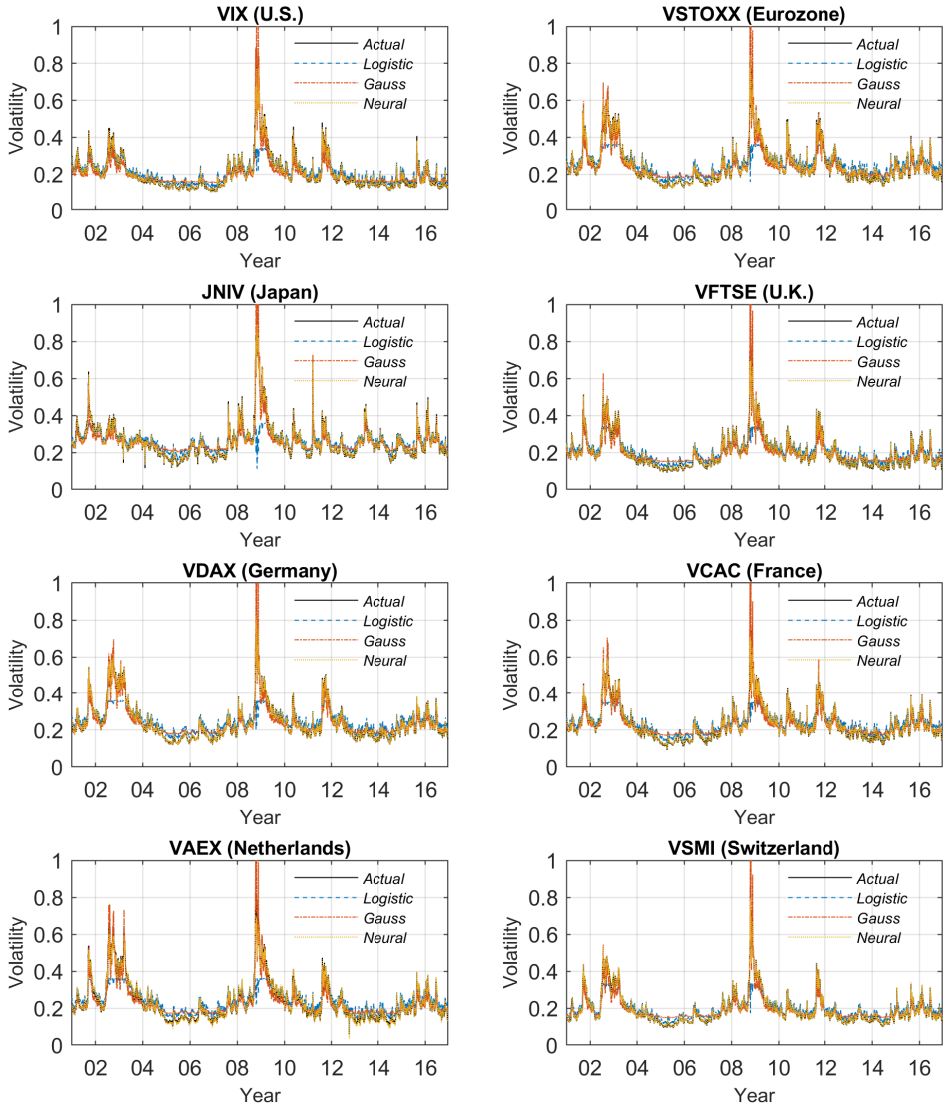


Figure 3: Actual *vs.* fitted chaotic volatility.

during the stock market downturn of 2002 and the global financial crisis of 2009. The Gaussian map, however, overestimates the risk during these turmoil periods. To formally investigate the goodness-of-fit of the estimated chaotic maps, we compute the general adjusted R-squared $\bar{R}^2$ of Magee (1990) for nonlinear models, the standard error of the estimate $S$, and the Bayesian information criterion (BIC) per observation in table 5. We further report the results of the optimal neural network function employed to estimate the Lyapunov exponent in the previous section 4.[8]

[8]The estimated coefficients of the neural net, the logistic, and the Gaussian maps are

Table 5: Goodness-of-fit of the chaotic maps.

| Volatility index | | VIX | VSTOXX | JNIV | VFTSE | VDAX | VCAC | VAEX | VSMI |
|---|---|---|---|---|---|---|---|---|---|
| Logistic map | $\bar{R}^2$ | 0.7178 | 0.6735 | 0.4732 | 0.7639 | 0.6703 | 0.7185 | 0.6757 | 0.7629 |
| | $S$ | 0.0471 | 0.0565 | 0.0661 | 0.0421 | 0.0571 | 0.0480 | 0.0608 | 0.0408 |
| | $\text{BIC}_o$ | -3.2697 | -2.9077 | -2.5922 | -3.4815 | -2.8866 | -3.2323 | -2.7602 | -3.5595 |
| Gaussian map | $\bar{R}^2$ | 0.8084 | 0.8562 | 0.8148 | 0.8234 | 0.8721 | 0.8478 | 0.8739 | 0.8319 |
| | $S$ | 0.0388 | 0.0375 | 0.0392 | 0.0367 | 0.0356 | 0.0353 | 0.0379 | 0.0343 |
| | $\text{BIC}_o$ | -3.6551 | -3.7263 | -3.6360 | -3.7699 | -3.8314 | -3.8453 | -3.7028 | -3.9017 |
| Neural net | $\bar{R}^2$ | 0.9651 | 0.9674 | 0.9561 | 0.9658 | 0.9753 | 0.9622 | 0.9702 | 0.9786 |
| | $S$ | 0.0166 | 0.0178 | 0.0191 | 0.0161 | 0.0156 | 0.0176 | 0.0184 | 0.0122 |
| | $\text{BIC}_o$ | -5.3482 | -5.2008 | -5.0403 | -5.4051 | -5.4664 | -5.2265 | -5.1355 | -5.9380 |

*Note:* This table reports the goodness-of-fit of the neural network, the logistic map, and the Gaussian map. Goodness-of-fit is evaluated with the general adjusted $\bar{R}^2$ (the highest the best) for nonlinear models (Magee, 1990), the standard error of the estimate $S$ (the lowest the best), and the Bayesian information criterion per observation $\text{BIC}_o$ (the lowest the best).

From table 5, the neural network, with orders reported in table 4, beats the other chaotic maps. Furthermore, the Gaussian map outperforms the logistic map with an $\bar{R}^2$ larger than 80% for all volatility indexes. An attractive characteristic of this map is that it has a simpler functional form, with merely two parameters, compared to the more complex neural net with a total of $q(m+2)+1$ estimated coefficients.

## 5.2 One-step-ahead volatility forecasts

We conduct out-of-sample volatility forecasting on an extended period starting from December 16, 2016 to December 31, 2017, with a total of 266 observations. The results of the one-day-ahead forecasting accuracy of the chaotic volatility models are evaluated with the Model Confidence Set (MCS) of Hansen et al. (2011). The MCS enables the comparison of a selection of competitive models under a specific loss function without knowledge of the true model. The MCS is performed using the mean squared-error (MSE) and the quasi-likelihood (QLIKE) loss functions as expressed in eqs. (13) and (14), respectively.[9]

available from the corresponding author upon request.

[9]In a related study, BenSaïda et al. (2018) compare the forecasting power of highly flexible stochastic models. Hence, a comparison between the predictive ability of chaotic maps and stochastic models becomes an interesting subject for a future research.

$$\text{MSE} = \frac{1}{T_f}\sum_{t=1}^{T_f}(v_t - \hat{x}_t)^2 \tag{13}$$

$$\text{QLIKE} = \frac{1}{T_f}\sum_{t=1}^{T_f}\left[\frac{v_t}{\hat{x}_t} - \ln\frac{v_t}{\hat{x}_t} - 1\right] \tag{14}$$

where $T_f$ is the number of forecasting data points, $v_t$ and $\hat{x}_t$ refer to the observed volatility and the variance forecast from a particular model, respectively.

The results are reported in table 6, and the one-step-ahead volatility forecasts from the chaotic maps are plotted in fig. 4. For both loss functions, MSE and QLIKE, the model confidence set shows that the neural network map outstandingly achieves the best performance for all markets, because this model has the smallest values of the loss functions among other forecasting models with $p$-values that equal to one. Surprisingly, the logistic map outperforms the Gaussian map in out-of-sample forecasting due to lower values of the loss functions, although it did not provide good in-sample fitting results.

The graphical inspection of fig. 4 demonstrates the remarkable forecasting power of the neural network for all data. Indeed, this function mimics the observed volatility pattern to a satisfying precision.

# 6 Conclusion

The presence of chaos in financial markets has been largely investigated in recent studies. However, the employed methods suffer from practical limitations related to the nature of the collected data. Indeed, the presence of noise due to market frictions, and the limited number of observations, which yields finite time series, are the two main drawbacks heavily discussed in the literature. Furthermore, chaos is barely investigated in the returns ignoring another determinant component to the dynamics of financial markets, which is the volatility.

In this paper, we present a robust method to test the existence of chaos in noisy and finite scalar time series. The approach is based on the estimation of the largest Lyapunov exponent using the Jacobian method, the derivation of the distribution of the exponent makes it consistent even in presence of moderate noise and finite data.

Applications are conducted on eight daily volatility indexes, relative to U.S., Eurozone, Japan, U.K., Germany, France, Netherlands, and Switzerland stock exchange markets. The results support the presence of low-level chaos in all investigated volatility indexes, with different complexity of the generating chaotic map. Consequently, to shed more light on the chaotic behavior of

Table 6: Volatility forecasting performance.

| Model | VIX | VSTOXX | JNIV | VFTSE | VDAX | VCAC | VAEX | VSMI |
|---|---|---|---|---|---|---|---|---|
| | MSE | | | | | | | |
| Logistic map | 0.6066 (0.000) | 1.2492 (0.000) | 1.1011 (0.000) | 0.5856 (0.000) | 1.1807 (0.000) | 1.0393 (0.000) | 1.1629 (0.000) | 0.5175 (0.000) |
| Gaussian map | 2.2571 (0.000) | 2.0393 (0.000) | 2.7790 (0.000) | 1.9996 (0.000) | 1.9599 (0.000) | 2.0232 (0.000) | 2.3725 (0.000) | 1.0551 (0.000) |
| Neural net | **0.0673** (1.000) | **0.1143** (1.000) | **0.0897** (1.000) | **0.0703** (1.000) | **0.0845** (1.000) | **0.2374** (1.000) | **0.1689** (1.000) | **0.0446** (1.000) |
| | QLIKE | | | | | | | |
| Logistic map | 0.0188 (0.000) | 0.0223 (0.000) | 0.0169 (0.000) | 0.0193 (0.000) | 0.0222 (0.000) | 0.0398 (0.019) | 0.0272 (0.000) | 0.0136 (0.000) |
| Gaussian map | 0.0593 (0.000) | 0.0362 (0.000) | 0.0395 (0.000) | 0.0565 (0.000) | 0.0362 (0.000) | 0.0528 (0.000) | 0.0524 (0.000) | 0.0269 (0.000) |
| Neural net | **0.0023** (1.000) | **0.0022** (1.000) | **0.0016** (1.000) | **0.0030** (1.000) | **0.0018** (1.000) | **0.0149** (1.000) | **0.0069** (1.000) | **0.0013** (1.000) |

*Note:* This table reports the mean losses of the different chaotic maps over the out-of-sample period (December 16, 2016 – December 31, 2017) with respect to two evaluation criteria (MSE $\times 10^3$ and QLIKE). Consistent $p$-values of the model confidence set (MCS) of Hansen et al. (2011) are reported (in parentheses) under the selected loss functions. The model with the highest $p$-value is given the best rank.

market risk, we model the volatility using different chaotic maps, namely, the logistic, the Gaussian, and the neural net. The findings show that neural networks are good fits to market volatility. Moreover, out-of-sample forecasting demonstrates the remarkable performance of neural nets into predicting the one-step-ahead volatility.

This study offers a better understanding to the nature of financial markets that can add to the existing literature. It also contributes to the conception of better models to predict the market risk considered as chaotic. Future research can investigate the efficiency of medium to long-term predictions of the neural networks when forecasting market volatility. Because deterministic chaotic systems have limited prediction horizon, which depends on the uncertainty of the initial conditions.

## Declarations of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

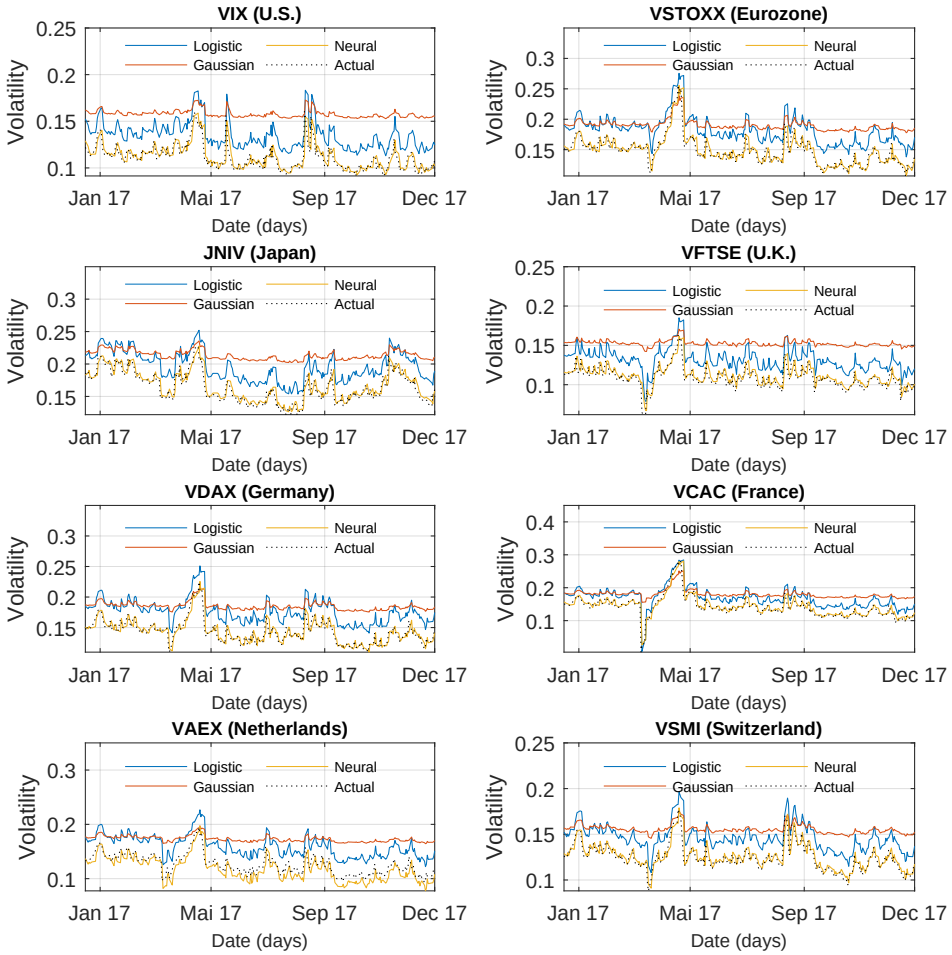


Figure 4: One-step-ahead chaotic volatility forecasts.

# References

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858.

Bahi, J. M., Couchot, J.-F., Guyeux, C., and Salomon, M. (2012). Neural networks and chaos: Construction, evaluation of chaotic networks, and prediction of chaos with multilayer feedforward networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(1):013122.

Bailey, B. A., Ellner, S., and Nychka, D. W. (1997). Chaos with confidence: Asymptotics and applications of local Lyapunov exponents. In Cutler, C. D. and Kaplan, D. T., editors, *Nonlinear Dynamics and Time Series: Building a Bridge Between the Natural and Statistical Sciences*, pages 115–133. Fields Institute Communications.

Barnett, W. A. and Hinich, M. J. (1992). Empirical chaotic dynamics in economics. *Annals of Operations Research*, 37(1):1–15.

BenSaïda, A. (2014). Noisy chaos in intraday financial data: Evidence from the American index. *Applied Mathematics and Computation*, 226:258–265.

BenSaïda, A. (2015). A practical test for noisy chaotic dynamics. *SoftwareX*, 3-4:1–5.

BenSaïda, A., Boubaker, S., Nguyen, D. K., and Slim, S. (2018). Value-at-risk under market shifts through highly flexible models. *Journal of Forecasting*, In Press.

BenSaïda, A. and Litimi, H. (2013). High level chaos in the exchange and index markets. *Chaos, Solitons & Fractals*, 54:90–95.

Brock, W., Dechert, D., Scheinkman, J., and LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, 15(3):197–235.

Brock, W. A. and Hommes, C. H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22(8):1235–1274.

Broomhead, D. and King, G. P. (1986). Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena*, 20(2):217–236.

Casdagli, M., Eubank, S., Farmer, J., and Gibson, J. (1991). State space reconstruction in the presence of noise. *Physica D: Nonlinear Phenomena*, 51(1-3):52–98.

Chan, J. C. C. and Grant, A. L. (2016). On the observed-data deviance information criterion for volatility modeling. *Journal of Financial Econometrics*, 14(4):772.

Coleman, T. F. and Li, Y. (1996). An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6(2):418–445.

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236.

Day, R. H. (1992). Complex economic dynamics: Obvious in history, generic in theory, elusive in data. *Journal of Applied Econometrics*, 7:S9–S23.

Dhanya, C. and Kumar, D. N. (2010). Nonlinear ensemble prediction of chaotic daily rainfall. *Advances in Water Resources*, 33(3):327–347.

Eckmann, J. P. and Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Review of Modern Physics*, 57:617–656.

Eraker, B., Johannes, M., and Polson, N. (2003). The impact of jumps in volatility and returns. *The Journal of Finance*, 58(3):1269–1300.

Faggini, M. (2014). Chaotic time series analysis in economics: Balance and perspectives. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 24(4):042101.

Farmer, J. D. and Sidorowich, J. J. (1987). Predicting chaotic time series. *Physical Review Letters*, 59:845–848.

Garcin, M. and Guégan, D. (2016). Wavelet shrinkage of a noisy dynamical system with non-linear noise impact. *Physica D: Nonlinear Phenomena*, 325:126–145.

Gottwald, G. A. and Melbourne, I. (2009). On the implementation of the $0-1$ test for chaos. *SIAM Journal on Applied Dynamical Systems*, 8(1):129–145.

Grassberger, P. and Procaccia, I. (1983a). Estimation of the Kolmogorov entropy from a chaotic signal. *Physical Review A*, 28:2591–2593.

Grassberger, P. and Procaccia, I. (1983b). Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9(1-2):189–208.

Guégan, D. (2009). Chaos in economics and finance. *Annual Reviews in Control*, 33(1):89–93.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, 79(2):453–497.

Hegger, R. and Kantz, H. (1999). Improved false nearest neighbor method to detect determinism in time series data. *Physical Review E*, 60:4970–4973.

Hilborn, R. (2000). *Chaos and nonlinear dynamics: an introduction for scientists and engineers*. Oxford University Press, 2 edition.

Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366.

Hu, J., Tung, W.-w., Gao, J., and Cao, Y. (2005). Reliability of the $0-1$ test for chaos. *Physical Review E*, 72:056207.

Kantz, H. (1994). A robust method to estimate the maximal Lyapunov exponent of a time series. *Physics Letters A*, 185(1):77–87.

Kennel, M. B., Brown, R., and Abarbanel, H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, 45:3403–3411.

Kim, H., Lee, K., Kyoung, M., Sivakumar, B., and Lee, E. (2009). Measuring nonlinear dependence in hydrologic time series. *Stochastic Environmental Research and Risk Assessment*, 23(7):907–916.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies*, 65(3):361–393.

Lai, D. and Chen, G. (2003). Distribution of the estimated Lyapunov exponents from noisy chaotic time series. *Journal of Time Series Analysis*, 24(6):705–720.

Li, X., Gao, G., Hu, T., Ma, H., and Li, T. (2014). Multiple time scales analysis of runoff series based on the chaos theory. *Desalination and Water Treatment*, 52(13-15):2741–2749.

Li, Y., Deng, S., and Xiao, D. (2011). A novel hash algorithm construction based on chaotic neural network. *Neural Computing and Applications*, 20(1):133–141.

Magee, L. (1990). $R^2$ measures based on Wald and likelihood ratio joint significance tests. *The American Statistician*, 44(3):250–253.

McCaffrey, D. F., Ellner, S., Gallant, A. R., and Nychka, D. W. (1992). Estimating the Lyapunov exponent of a chaotic system with nonparametric regression. *Journal of the American Statistical Association*, 87(419):682–695.

Nakajima, J. and West, M. (2012). Dynamic factor volatility modeling: A bayesian latent threshold approach. *Journal of Financial Econometrics*, 11(1):116.

Nychka, D., Ellner, S., McCaffrey, D., and Gallant, A. R. (1992). Finding chaos in noisy systems. *Journal of the Royal Statistical Society. Series B*, 54:399–426.

Pal, S. K. and Mitra, S. (1992). Multilayer perceptron, fuzzy sets, and classification. *IEEE Transactions on Neural Networks*, 3(5):683–697.

Rubežić, V., Djurović, I., and Sejdić, E. (2017). Average wavelet coefficient-based detection of chaos in oscillatory circuits. *COMPEL - The international journal for computation and mathematics in electrical and electronic engineering*, 36(1):188–201.

Shintani, M. and Linton, O. (2004). Nonparametric neural network estimation of Lyapunov exponents and a direct test for chaos. *Journal of Econometrics*, 120(1):1–33.

Takens, F. (1981). Detecting strange attractors in turbulence. In Rand, D. and Young, L.-S., editors, *Dynamical Systems and Turbulence, Warwick 1980*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381. Springer Berlin Heidelberg.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena*, 58(1):77–94.

Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. (1985). Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(3):285–317.

Yang, Q., Zeng, C., and Wang, C. (2013). Fractional noise destroys or induces a stochastic bifurcation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(4):043120.