Research article

Analysis of chaotic economic models through Koopman operators, EDMD, Takens' theorem and Machine Learning

John Leventides, Evangelos Melas, Costas Poulios[*] **and Paraskevi Boufounou**

Department of Economics, Faculty of Economics and Political Sciences, National and Kapodistrian University of Athens, 1, Sofokleous str. 10559, Athens, Greece

* **Correspondence:** konpou@econ.uoa.gr; Tel: +302103689412; Fax: +302103689400.

Abstract: We consider dynamical systems that have emerged in financial studies and exhibit chaotic behaviour. The main purpose is to develop a data-based method for reconstruction of the trajectories of these systems. This methodology can then be used for prediction and control and it can also be utilized even if the dynamics of the system are unknown. To this end, we combine merits from Koopman operator theory, Extended Dynamic Mode Decomposition and Takens' embedding theorem. The result is a linear autoregressive model whose trajectories approximate the orbits of the original system. Finally, we enrich this method with machine learning techniques that can be used to train the autoregressive model.

1. Introduction

Chaotic dynamical systems are deterministic dynamical systems which are characterized by the fact that small changes in the initial conditions lead to very large differences at a later time. Chaotic dynamical systems are to be contradistinguished with stochastic dynamical systems. Stochastic dynamical systems are not deterministic and a probability space is required to study their evolution.

In recent years, chaotic systems have been receiving more and more attention and interest because of their potential applications in many areas which range from physics and telecommunications to biological networks and economic models. Since the chaotic phenomenon in economics was discovered in 1985, it has been widely accepted that the economy and the finance systems are very complicated nonlinear systems containing many complex factors and accordingly, many efforts have been devoted to their study (Chian et al., 2006; Gao and Ma, 2009; Gu´egan, 2009; Wijeratne et al., 2009). As stated in

Chian et al. (2006), characterization of the complex dynamics of economic systems is a powerful tool for pattern recognition and forecasting of business and financial cycles, as well as for optimization of management strategy and decision technology. Therefore, the need for analysis and control of these systems is emerging and many researchers have been working towards this direction.

In this paper, we indicate a method for approximating a macroeconomic chaotic system and reconstructing its trajectories using a linear finite dimensional dynamical system. This method sits across Koopman operators, EDMD, Takens theorem and Machine Learning. EDMD and its precursor DMD, as well as, Koopman operators, Takens' theorem and Machine Learning have been extensively used in Finance (for instance, Hua et al., 2017; Mauroy et al., 2020; Mann and Kutz, 2016; Stavroglou et al., 2019; Ni et al., 2021).

Roughly speaking, the Koopman operator "lifts" the dynamics of the original system from the state space to linear spaces consisting of functions defined on the state space (which are called observables). The advantage of this lifting is that we obtain a linearization of the original system which holds to the whole state space and many properties of the dynamics correspond to spectral properties of the Koopman operator. The disadvantage is that the linear system induced by the Koopman operator is infinite dimensional. Therefore, its dynamics and spectral properties cannot be calculated while several methods for analysis and control cannot be applied.

The Extended Dynamic Mode Decomposition provides us with a method for obtaining finite dimensional approximations of the Koopman operator. These approximations, under some suitable assumptions, converge to the Koopman operator. Therefore, they can also be used to reconstruct the trajectories in the phase space of the original nonlinear system.

The success of EDMD depends on a set of observables (a *dictionary*) which is chosen a priori. This method is purely data-driven and is based on instantaneous measurements of the values of the observables. Then, the linear dynamical system that is provided advances the measurements form one time to the next.

On the other hand, it is generally feasible to extract information about features of phase space from time series of general measurements made on an evolving system (Sauer et al., 1991). The central result in this direction was proved by (Takens, 1981) who showed how a time series of measurements of a single observable can be used to reconstruct qualitative features of the phase space of the system. The technique described by Takens, the *method of delays*, is so simple that it can be applied to essentially *any* time series whatever, and has made possible the wide−ranging search for chaos in dynamical systems. Takens' theorem asserts that the phase space of a dynamical system and in particular strange attractors of the system can be diffeomorphically embedded in a higher dimensional space spanned by an appropriate number of time lags of a single observable of the system.

This is a very powerful result since it allows to recover the dynamics from the observation of just *one* variable, the embedding preserves the topology of the phase space and it preserves the topology of the attractor in particular, the embedding being a diffeomorphism allows system identification, it also allows to calculate the box-counting dimension of the attractor, and it allows to calculate the (positive) Lyapunov exponents. However, Takens' theorem does not allow reconstruction of the trajectories of the dynamical system. This is to be juxtaposed with the Koopman-EDMD theory which aims at reconstructing the trajectories of the system.

What the two approaches have in common is that both are data-driven since the observables they use are constructed from data measured during the evolution of the system. In Takens' theorem the observables are simpler in form and they are the time lags of a single observable. In Koopman-EDMD theory the observables are usually the state variables of the dynamical system and non−linear functions of them. In this paper we use an hybrid method in order to reconstruct with simple means trajectories of a macroeconomic dynamical system which exhibits chaotic behaviour. In particular we use Koopman-EDMD theory with observables those used by Takens' Theorem. The finite dimensional approximation to the Koopman operator which drives the dynamics and reconstructs the trajectories of the system is obtained by two different methods: By a Linear autoregressive model and by an Autoregression with machine learning methods. The use of the Linear autoregressive model leads to more faithful reconstruction of the trajectories of the dynamical system at hand than the use of Autoregression with machine learning methods. However, this is attributed to the fact the error made in the case of the Linear autoregressive model can be corrected whereas this is not feasible in the case of Autoregression with machine learning methods where the error made in the prediction of the future value of the observable from the previous values cannot be corrected and accumulates over time.

Therefore, the motivation, innovation and contribution of this paper can be summarized as follows: This paper addresses the trajectory approximation of an hyperchaotic system via EDMD methods. The EDMD methods give rise to a linear system on an enhanced state space that can approximate a given trajectory. Having data of a given trajectory in a finite horizon allows to construct a discrete linear system of dimension $n >> m$, where m is the dimension of the state space of the nonlinear system. Takens' Theorem indicates how large n can be. Here we demonstrate the approximation of a single trajectory of a chaotic system via EDMD methods equivalent to linear Autoregressive construction. Furthermore, we demonstrate the use of the same method on multiple trajectories. Finally, we demonstrate how we can use the same data in order to construct nonlinear trajectory approximation via machine learning methods.

Trajectory approximation of chaotic nonlinear systems is a particularly difficult problem as these trajectories are sensitive to initial conditions and also exhibit variable spectrum and almost periodicity properties. Simultaneous approximation of orbits that differ seems a challenge via any method requiring big data and accurate approaches. The use of linear dynamic models seems in certain respects more appropriate as the knowledge of its structure allows to correct the fitted model if significant part of the spectrum of the data is lost due to numerical or other errors. This is demonstrated in this paper. The conclusions in this paper are derived with the use of an economic model (Yu et al., 2012) which exhibits hyperchaotic behaviour. We note that the scope the methods and the results in this paper are entirely different from those in Yu et al. (2012). In a nutshell in this paper we approximate the trajectory of the hyperchaotic system introduced in Yu et al. (2012) via EDMD methods whereas in Yu et al. (2012) the general structure of this system is studied, namely, its equilibrium points, its bifurcations, e.t.c..

The EDMD method is data driven. Consequently our method can be applied to any dynamical system for which probably the dynamical law which drives its dynamics is unknown and data, in the form of time series, can be collected for some of its trajectories in state space. Moreover, our approach can be advocated for any nonlinear dynamical system whose dynamical law which drives its dynamics might be known but a linearization of its dynamics via the EDMD method may be required in order for example to study control theory of this linearized system; control theory of linear systems is much better understood than the control theory of nonlinear systems.

The rest of the paper is organized as follows. In Section 2, we firstly review a dynamical system which has emerged from financial studies and it shows chaotic behaviour for some values of the parameters. Furthermore, we briefly present some basic facts about the Koopman operator. We also describe the EDMD algorithm and the main idea about the connections with Takens' theorem. The latter is analyzed with more details in Subsection 2.5. In Section 3 we consider the chaotic system of Section 2.1 and we developed our methodology for approximating and reconstructing a single trajectory of the system. In Section 4, we demonstrate the proposed approach in the case where we wish to simultaneously approximate two orbits of the original system with the same autoregressive model. Finally, Section 5 concludes the paper.

2. Preliminaries

2.1. Chaotic and hyperchaotic macroeconomic models

Various financial system have been found to exhibit chaotic dynamics (Chian, 2000; Fanti and Manfredi, 2007; Gu´egan, 2009; Hass, 1998; Holyst et al., 1996; Lorenz, 1993). A lot of research has been invested in the analysis, control and stabilization of these systems (for instance, see Ahmad et al., 2021; Chen et al., 2021; Ma et al., 2022; Pan and Wu, 2022; for some recent studies). In the present paper, we use in our analysis a chaotic dynamical model that describes the time evolution of three state variables: x is the interest rate, y is the investment demand and z is the price exponent. For more details about the origination, the development and the economic interpretation of the model, we refer to Ma and Chen (2001), Yu et al. (2012), Jian et al. (2009) and Zhao et al. (2011). The system is described by the following non-linear differential equations.

$$\begin{cases} \dot{x} = & z + (y - a)x \\ \dot{y} = & 1 - by - x^2 \\ \dot{z} = & -x - cz \end{cases} \tag{1}$$

where a, b, c are positive constant parameters that represent the amount of savings, the cost per investment and the elasticity of demand respectively (see Ma and Chen, 2001).

When the values of parameters are chosen as $a = 0.9$, $b = 0.2$, $c = 1.2$ and for initial conditions $(x_0, y_0, z_0) = (1, 3, 2)$ the system (1) has a strange attractor that can be classified neither as a stable equilibrium nor as a periodic or almost periodic oscillation. Therefore, the system exhibits chaotic behaviour (see Ma and Chen, 2001; Rigatos, 2017).

Later, in Yu et al. (2012), it was observed that x depends not only on the investment demand and the price exponent, but also on the average profit margin (denoted by $u(t)$). Therefore, in Yu et al. (2012) a dynamical system consisting of four state variables (x, y, z, u) is proposed which is described by the next first order non-linear differential equations

$$\begin{cases} \dot{x} = & z + (y - a)x + u \\ \dot{y} = & 1 - by - x^2 \\ \dot{z} = & -x - cz \\ \dot{u} = & -dxy - ku \end{cases}$$

where a, b, c, d, k are positive parameters. It is shown that, if the parameters' values are chosen to be $a = 0.9$, $b = 0.2$, $c = 1.5$, $d = 0.2$, $k = 0.17$, then the system shows *hyperchaotic* behaviour (see Yu et

al., 2012). A hyperchaotic system is usually defined as a chaotic system with more than one positive Lyapunov exponent. Therefore, these system show more complex behaviour than chaotic systems.

2.2. Koopman operators

The Koopman operator theory represents an increasingly popular formalism of dynamical systems since it allows the analysis of nonlinear and complicated systems. Therefore, it can be utilized in chaotic and hyperchaotic systems.

Assume that $(\mathcal{M}, f)$ is a continuous dynamical system, where $\mathcal{M} \subseteq \mathbb{R}^n$ is a real manifold, f is the evolution map and the system is described by the equation $\dot{\mathbf{x}} = f(\mathbf{x})$. We also denote by $\Phi_t(\mathbf{x}_0)$ the flow map defined as the state of the system at time t when we start from the initial condition $\mathbf{x}_0$.

Any function $g \colon \mathcal{M} \to \mathbb{C}$ is called an *observable* of the system $(\mathcal{M}, f)$. Let now $\mathcal{F}$ be a function space of observables which enjoys the property: for any $g \in \mathcal{F}$ and any $t \geq 0$, it holds that $g \circ \Phi_t \in \mathcal{F}$. Then, the Koopman operator, which is actually a semigroup of operators, is defined for any $t \geq 0$, as $K_t \colon \mathcal{F} \to \mathcal{F}$ with

$$K_t(g) = g \circ \Phi_t.$$

It is easily verified that K_t is linear for any $t \geq 0$. Throughout this paper, following the common practice, we will use the term Koopman operator to refer to the whole class of operators $\mathcal{K} = (K_t)_{t \geq 0}$. In many applications, the space $\mathcal{F}$ coincides with the Hilbert space $L^2(\mathcal{M})$, however other choices can be used as well.

Similarly, the Koopman operator can be defined for any discrete dynamical system $x_{k+1} = f(x_k)$ defined on the state space $\mathcal{M}$. The Koopman operator is the composition of any observable with the evolution map f, that is, $\mathcal{K} \colon \mathcal{F} \to \mathcal{F}$ is defined by $\mathcal{K}(g) = g \circ f$, for any observable $g \in \mathcal{F}$. Here, $\mathcal{F}$ is a linear space of observables which is closed under composition with f and the space $L^2(\mathcal{M})$ remains a standard choice.

Roughly speaking the Koopman operator updates every observable in the space $\mathcal{F}$ according to the evolution of the dynamical system. It defines a linear dynamical system $(\mathcal{F}, \mathcal{K})$ which completely describes the original system $(\mathcal{M}, f)$. This linearization is applied to any nonlinear system and it is global, i.e. it does not hold only to the area of some attractor or fixed point. Furthermore, many properties of the original system can be codified as properties of the Koopman operator and, usually they can be related to the eigenstructure of $\mathcal{K}$. However, since $\mathcal{K}$ is usually infinite dimensional, its spectral properties cannot be calculated except some rare cases.

To overcome the problem of dimensionality, one has to look for finite dimensional approximations of the Koopman operator, i.e. finite dimensional subspaces of the domain of $\mathcal{K}$ which are invariant under $\mathcal{K}$. The most natural candidates for this project are the spaces generated by eigenvectors of $\mathcal{K}$, which, however, are difficult to find. Hence, we are obliged to search for finite dimensional linear approximations of the Koopman operator. Towards this direction, the Dynamic Mode Decomposition (DMD) and its generalization the so-called Extended Dynamic Mode Decomposition (EDMD) have been proved very successful.

2.3. Extended Dynamic Mode Decomposition (EDMD)

We next describe briefly the main steps of the EDMD algorithm. Firstly, one has to consider a set of observables $\{g_1, g_2, \ldots, g_p\}$. This set is usually called a *dictionary*. The cardinality p of the dictionary is expected to be much bigger than n (the dimension of the original state space $\mathcal{M}$). The DMD (Dynamic

Mode Decomposition) uses only the observables $g_i(\mathbf{x}) = x_i$, for $i = 1, 2, \ldots, n$. However, in EDMD we can choose any observables. Hence, an augmented state space is constructed and EDMD provides, in general, better approximations than DMD. We denote the augmented state space by $\overline{\mathcal{M}}$ and its elements by $\mathbf{y} = [g_1(\mathbf{x}), \ldots, g_p(\mathbf{x})]^T$.

The second step is to collect the data. We consider a trajectory of the original system with initial condition $\mathbf{x}_0$. The trajectory is witnessed for some finite time horizon T and sampling points are collected at a fixed time interval ΔT. Uniform sampling is not mandatory and other sampling methods can be applied. Therefore, $n_0 = \frac{T}{\Delta T}$ points in this trajectory are considered. We denote our data by

$$(\mathbf{x}_s)_{s=0}^{n_0}.$$

These data correspond to data $(\mathbf{y}_s)_{s=0}^{n_0}$ in the augmented space $\overline{\mathcal{M}}$. We now consider the data matrices

$$\mathbf{Y}_{[0,n_0-1]} = [\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{n_0-1}] \quad \text{and} \quad \mathbf{Y}_{[1,n_0]} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{n_0}].$$

Finally, using least square regression methods, we obtain a $p \times p$ matrix $\mathbf{A}$, such that $\mathbf{Y}_{[1,n_0]} \sim \mathbf{A}\mathbf{Y}_{[0,n_0-1]}$. Therefore, we have

$$\mathbf{A} = \operatorname*{argmin}_{\tilde{\mathbf{A}} \in \mathbb{R}^{p \times p}} \left\| \mathbf{Y}_{[1,n_0]} - \tilde{\mathbf{A}}\mathbf{Y}_{[0,n_0-1]} \right\|.$$

Here, $\| \cdot \|$ denotes some matrix norm, for instance the Frobenius norm.

This procedure can also be applied to several trajectories. In this case, we construct data matrices $\mathbf{Y}_{j[0,n_0-1]}$ and $\mathbf{Y}_{j[01,n_0]}$ for every trajectory $j = 1, 2, \ldots, k$. Then, the matrix $\mathbf{A}$ is chosen such that

$$\mathbf{A} = \operatorname*{argmin}_{\tilde{\mathbf{A}} \in \mathbb{R}^{p \times p}} \sum_{j=1}^{k} \left\| \mathbf{Y}_{j[1,n_0]} - \tilde{\mathbf{A}}\mathbf{Y}_{j[0,n_0-1]} \right\|.$$

Roughly speaking, $\mathbf{A}$ is a best-fit matrix which relates the two data matrices in every trajectory. This matrix generates a finite dimensional linear system that advances spatial measurements from one time to the next and therefore it may provide approximations of the original nonlinear system.

One of the advantages of the EDMD method is that it depends on data. Hence, it may be applied when the dynamics of the system is unknown. On the other hand, the success of this method depends on the dictionary which is chosen a priori. In many cases, it is a difficult problem to find the suitable observables. Recent studies try to utilize machine learning and artificial intelligence methods in order to "train" the dictionary.

2.4. Koopman operator and time-delayed coordinates

As it has been mentioned, EDMD advances measurements of the state of the system from one time to the next. However, we may also obtain measurements coordinates for the approximation of Koopman operator using time-delayed measurements of the system. This approach is also data-driven and utilize the information from the previous measurements to predict the future. It is particularly useful in chaotic systems. In the case of fixed points or periodic orbits, more data from previous measurements have small contributions. On the contrary, when the trajectories densely fill a strange attractor, more data provide more information.

From this point of view, Koopman operators can interact with Takens embedding theorem. The latter could provide us information about the dimensions of the augmented state-space as well as the dictionary that should be chosen. A connection between Koopman operators and Takens Theorem has already been investigated in Mezić and Banaszuk (2004).

2.5. Takens' theorem

We firstly need some definitions. A dynamical system on a manifold M is specified by a flow

$$\phi_{(\cdot)}(\cdot) : \mathbb{R} \times M \to M,$$

where $\phi_0(x) = x$, ϕ_t is a diffeomorphism for all t and $\phi_t \circ \phi_s = \phi_{t+s}$ (e.g. ϕ is the solution to a system of ODEs).

Roughly speaking, an attractor of a dynamical system is a subset of the state space to which orbits originating from typical initial conditions tend as time increases. It is very common for dynamical systems to have more than one attractor. For each such attractor, its basin of attraction is the set of initial conditions leading to long−time behavior that approaches that attractor.

An **attracting set** A for the flow ϕ_t is a closed set such that:

- the basin of attraction $B(A)$ has positive measure,
- for $A' \subseteq A$ the difference $B(A) \setminus B(A')$ has positive measure.

A set $A \subset M$ is an **attractor** if it is an attracting set that contains a dense orbit of the flow ϕ_t. An attractor A is **strange attractor** ((Eckmann and Ruelle, 1985; Guckenheimer and Holmes, 1983), if its box-counting dimension d is non-integer.

Theorem 2.1 (Takens' theorem). *Let M be a compact manifold of (integer) dimension q. Then for generic pairs (ϕ, y), where*

- *$\phi : M \to M$ is a C^2−diffeomorphism of M in itself,*
- *$y : M \to M$ is a a C^2−differentiable function,*

the map $\Phi_{(\phi, y)} : M \to \mathbb{R}^{2q+1}$ given by

$$\Phi_{(\phi, y)}(x) := \left(y(x), y(\phi(x)), y(\phi^2(x)), ..., y(\phi^{2q}(x)) \right)$$

is an embedding (i.e. an injective and an immersive map) of M in $\mathbb{R}^{2q+1}$.

If ϕ_t is a flow on M and $\tau > 0$ is a fixed time, then we can define the delay map by

$$\Phi(x) := (y(x), y(\phi_{-\tau}(x)), y(\phi_{-2\tau}(x)), ..., y(\phi_{-\kappa\tau}(x))). \tag{2}$$

If M is a manifold that is an attractor for the flow ϕ_t, then $\phi_{-\tau}$ is a diffeomorphism of M into itself, so if $\kappa \geq 2\dim(M) + 1$ for generic y, τ, Takens' theorem (see Sauer et al., 1991; Packard et al., 1980; Takens, 1981; Muldoon et al., 1993) says that the delay map defined by (2) is actually an embedding.

Finally, we close with some details about the meaning of generic. Let P be a property of functions in $C^k(M, N)$ they might or might not have. Then we say that $P(f)$ is true for generic $f \in C^k(M, N)$ if the set of functions for which it holds is open and dense in the C^k topology. So arbitrary small perturbations turn bad choices in good choices.

The central plank of this theory is a result– suggested by several people (Packard et al., 1980) and eventually proved in (Takens, 1981)– which shows how a time series of measurements of a single observable can often be used to reconstruct qualitative features of the phase space of the system. The technique described by Takens, the *method of delays* (2), is so simple that it can be applied to essentially any time series whatever, and has made possible the wide-ranging search for chaos.

3. Application to chaotic economic models

We consider the 4-dimensional (financial) dynamical system described in Section 2.1, which for the parameter values $a = 0.9$, $b = 0.2$, $c = 1.5$, $d = 0.2$, $k = 0.17$ exhibits chaotic behaviour. We select initial conditions for the four variables to be: $x(0) = 1$, $y(0) = 5$, $z(0) = 1$, $u(0) = 1$. For this values the trajectory for $x(t)$ is depicted in Figure 1.
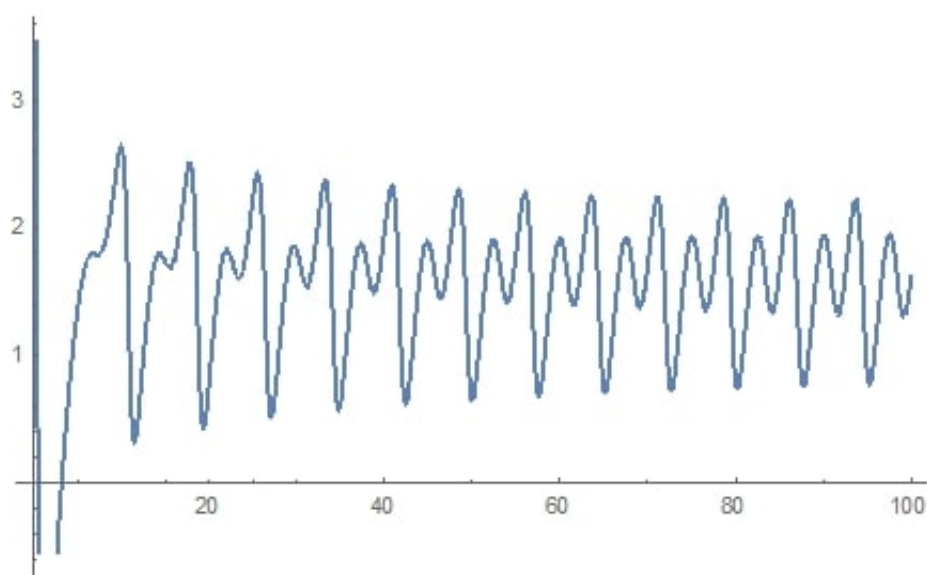
Figure 1. The trajectory of the investment rate $x(t)$, for initial values $x(0) = 1$, $y(0) = 5$, $z(0) = 1$, $u(0) = 1$

The vector of variables (x, y, z, u) converges to a chaotic attractor in $\mathbb{R}^4$. According to Whitney Takens theorem we can embed this attractor to a higher dimensional space of dimension d at least 8 by considering a single observable and a vector of d of its lagged values. We try to use this idea so that to attempt to reconstruct the above trajectory by using a 9-dimensional dynamical system. The simplest way to do that is to consider x as an observable and to model the evolution of the vector $(x(T - 9\Delta t), \ldots, x(T))$ assuming a sampling period Δt in our case $\Delta t = 1$.

3.1. Linear autoregressive model

The simplest model for this evolution is a linear autoregressive model

$$x_{n+9} = a_8 x_{n+8} + \ldots + a_0 x_n.$$

We model the time series of the above plotted variable $x(t)$ obtained after sampling and we get the results showed in Tables 1 and 2. Using these results, we reconstruct the x trajectory as shown in Figure 2.

Table 1. Basic parameter table. We have adjusted $R^2 = 0.986$.

	Estimate	Standard error	t-Statistic	p-value
♯1	0.382345	0.0526912	7.25634	1.1104×10^{-11}
♯2	0.0528833	0.0593214	0.891471	0.373854
♯3	0.0510388	0.0532897	0.95776	0.339454
♯4	0.0669098	0.0533061	1.2552	0.211016
♯5	−0.0546443	0.0532567	−1.02606	0.306227
♯6	−0.0756992	0.0534338	−1.41669	0.158283
♯7	0.488672	0.0536388	9.11044	1.47733×10^{-16}
♯8	0.457925	0.0622202	7.35974	6.12628×10^{-12}
♯9	−0.374469	0.040841	−9.16895	1.01965×10^{-16}

Table 2. Analysis of variance.

	DF	SS	MS	F-Statistic	p-value
♯1	1	440.283	440.283	12718.8	2.36926×10^{-170}
♯2	1	0.252312	0.252312	7.28877	0.00759293
♯3	1	39.9561	39.9561	1154.25	1.03165×10^{-80}
♯4	1	0.388756	0.388756	11.2304	0.000978538
♯5	1	0.447058	0.447058	12.9146	0.000419584
♯6	1	4.09436	4.09436	118.277	1.52056×10^{-21}
♯7	1	9.81632	9.81632	283.573	5.71589×10^{-39}
♯8	1	0.0967607	0.0967607	2.79521	0.0962645
♯9	1	2.9102	2.9102	84.0696	1.01965×10^{-16}
Error	182	6.30022	0.0346166		
Total	191	504.545			
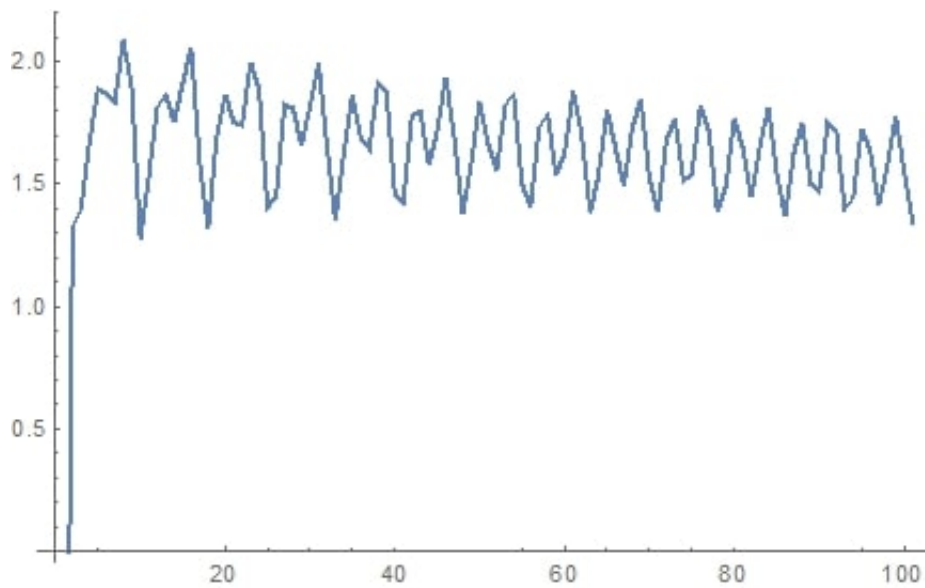
Figure 2. Reconstruction of the trajectory of *x*.

Furthermore, we extrapolate the modes of the above AR system which have modulus close to unity on the unit circle by radial projection. The new AR system reconstructs $x(t)$ as depicted in Figure 3.
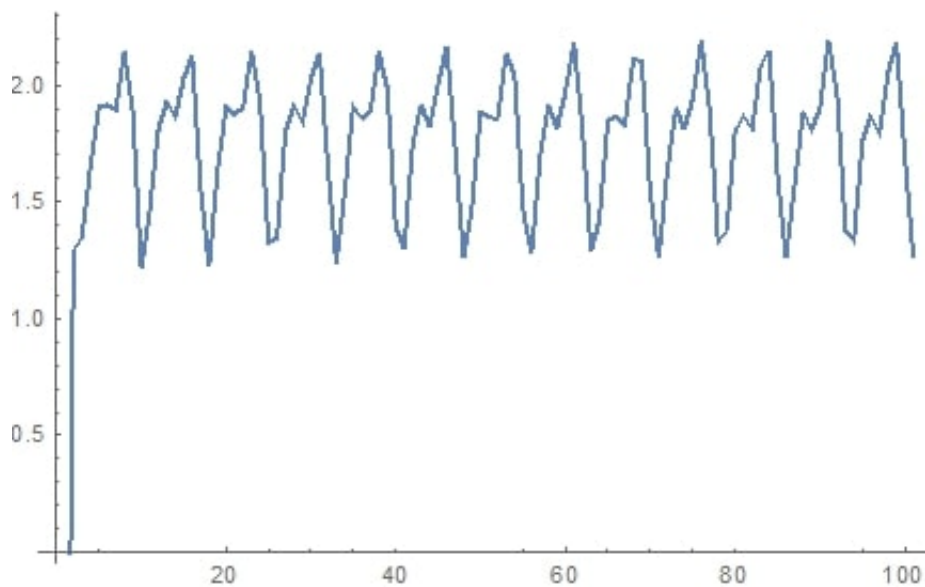
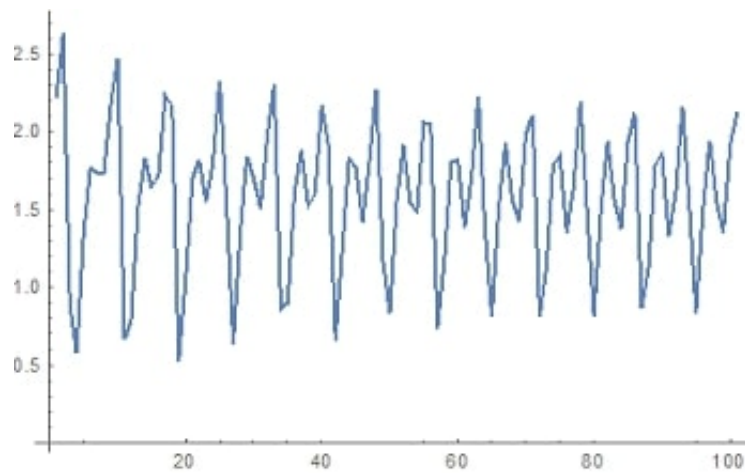**Figure 3.** Reconstruction of the trajectory of *x*.

The above results demonstrate that the considered chaotic financial system can be approximated, at least orbitwise and in finite horizon, by AR models of high enough dimensions and thus EDMD methods may assist in modeling the Takens embedding for the purpose of simulating single or multiple chaotic trajectories.

### 3.2. Autoregression with machine learning methods

Finally, we apply two further methods to explore nonlinear relationships for the transition proposed by the Takens embedding using machine learning methods. Both are generally described by a nonlinear autoregression

$$x_{n+9} = F(x_{n+8}, \ldots, x_n),$$

which was realised by two machine learning approaches, namely random forest interpolation and Gaussian process interpolation. The results for the same trajectory reconstruction are depicted in Figures 4 (for random forests) and 5 (for Gaussian process interpolation). Details of the two approaches are given in Tables 3 and 4.

Figure 4. Reconstruction of the trajectory of *x* using random forests.
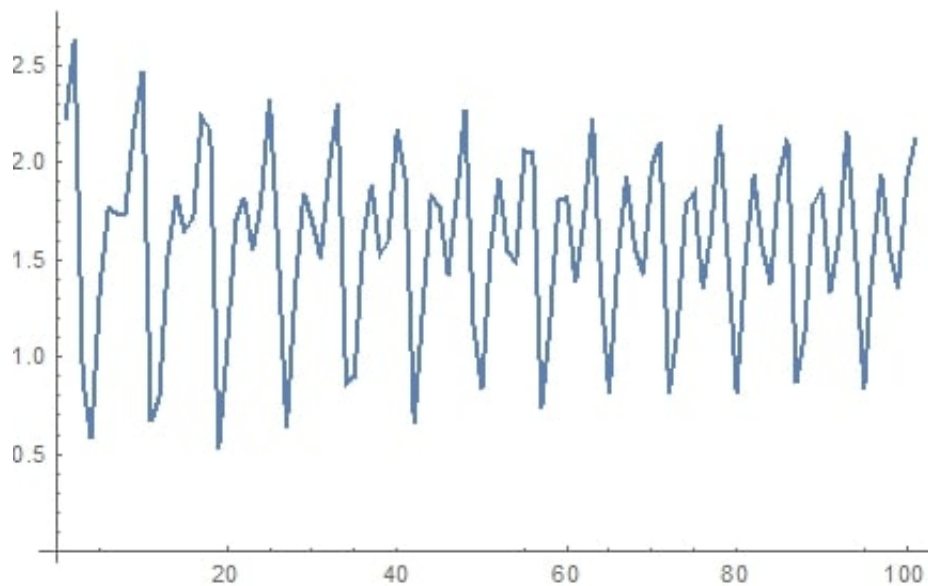
Figure 5. Reconstruction of the trajectory of *x* using Gaussian interpolation process.

Table 3. Information for the random forest method.

Predictor information	
Method	Random Forest
Number of features	1
Number of training examples	190
Number of trees	50

**Table 4.** Information for the Gaussian interpolation process.

Predictor information	
Method	Gaussian process
Number of features	1
Number of training examples	190
Assume Deterministic	False
Numerical Covariance Type	Squared Exponential
Nominal Covariance Type	Hamming Distance
Estimation Method	Maximum Posterior
Optimization Method	Find Minimum

Both methods fit well to the training data however when used for trajectory reconstruction starting from the same initial condition the subsequent calculation of the trajectory suffers from error accumulation and therefore reconstruction worsens with time. Gaussian process interpolation seems to work better.

4. Simultaneous reconstruction of two orbits

In this section, we try to simultaneously approximate two orbits of the hyperchaotic dynamical system via a single autoregressive dynamical system. To achieve this goal we need to increase the number of lags to 30.

First of all, we use data from numerical integration to plot the second orbit. This orbit is shown in Figure 6 and corresponds to initial conditions $x(0) = 0$, $y(0) = 2$, $z(0) = 0$ and $u(0) = 1$ and time horizon $t \in [0, 100]$. We note that the new initial conditions give rise to an orbit that is quite far from the first one. Hence, the task of the simultaneous approximation is quite challenging. The new autoregression is supposed to produce a new dynamical system that approximate the two trajectories by a single autoregressive equation and different initial conditions.
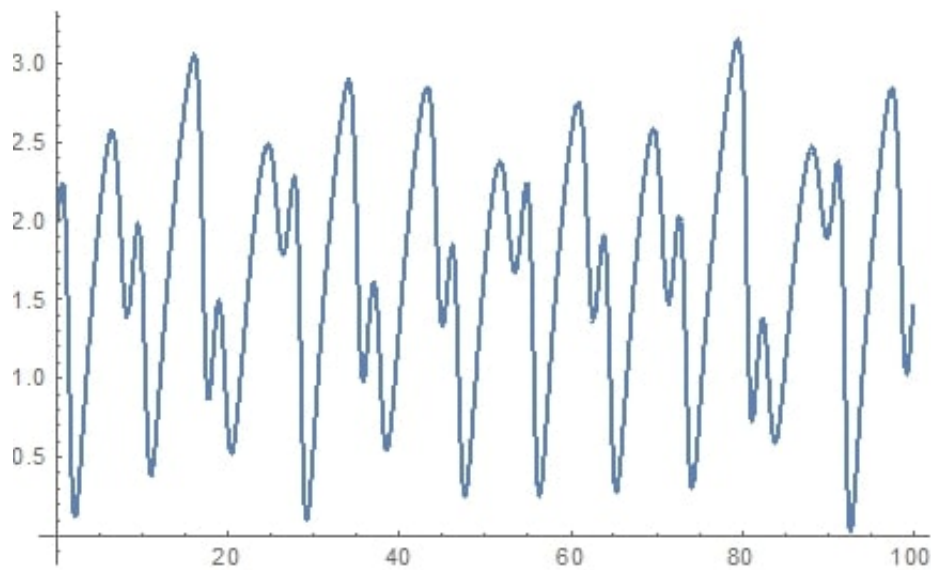
Figure 6. The trajectory of the investment rate $x(t)$, for initial values $x(0) = 0$, $y(0) = 2$, $z(0) = 0$, $u(0) = 1$ and time horizon $T = 100$.

4.1. Linear autoregressive model.

Similarly to the case of one trajectory, the simplest autoregressive model that one may consider is linear. The regression results are depicted in Table 6. The value of the adjusted R^2 is 0.9777954946531674. We can see that this value is smaller than the corresponding value of the one orbit case, however, it is still very close to 1. The regression function is given by:

$$x_{n+31} = \sum_{j=1}^{30} a_j x_{n+j},$$

where $(a_j)_{j=1}^{3} 0$ are the estimates given by the first column in Figure 6.

Finally, we can use the above linear autoregressive dynamical system to approximate the two different original trajectories of $x(t)$. Figure 7 shows the plots of these approximations.

Table 5. Analysis of variance.

	DF	SS	MS	F-Statistic	p-value
♯1	1	885.406	885.406	13286.5	1.38351×10^{-256}
♯2	1	0.466206	0.466206	6.99591	0.00858584
♯3	1	24.5391	24.5391	368.235	1.2149×10^{-54}
♯4	1	1.85537	1.85537	27.8417	2.47799×10^{-7}
♯5	1	2.68491	2.68491	40.2899	7.76562×10^{-10}
♯6	1	17.6117	17.6117	264.283	2.08632×10^{-43}
♯7	1	23.8109	23.8109	357.307	1.51675×10^{-53}
♯8	1	11.7662	11.7662	176.565	3.37191×10^{-32}
♯9	1	6.70749	6.70749	100.653	1.07024×10^{-20}
♯10	1	0.204543	0.204543	3.06939	0.0807683
♯11	1	8.20769	8.20769	123.165	2.54294×10^{-24}
♯12	1	0.0414977	0.0414977	0.622716	0.430643
♯13	1	0.0341088	0.0341088	0.511839	0.474883
♯14	1	1.21188	1.21188	18.1856	0.0000266385
♯15	1	2.18295	2.18295	32.7575	2.46335×10^{-8}
♯16	1	1.68824	1.68824	25.3338	8.1894×10^{-7}
♯17	1	0.371216	0.371216	5.57049	0.0188848
♯18	1	4.50535	4.50535	67.6076	5.53381×10^{-15}
♯19	1	0.313024	0.313024	4.69725	0.030972
♯20	1	0.360659	0.360659	5.41207	0.0206423
♯21	1	0.718296	0.718296	10.7788	0.00114403
♯22	1	1.03007	1.03007	15.4573	0.000104143
♯23	1	1.23863	1.23863	18.5869	0.0000218361
♯24	1	0.0496161	0.0496161	0.744542	0.388877
♯25	1	0.00615967	0.00615967	0.0924324	0.761311
♯26	1	0.202932	0.202932	3.0452	0.0819668
♯27	1	2.1012	2.1012	31.5307	4.36478×10^{-8}
♯28	1	0.157497	0.157497	2.36341	0.125231
♯29	1	0.199552	0.199552	2.99448	0.0845438
♯30	1	0.0705909	0.0705909	1.05929	0.304179
Error	310	20.6583	0.0666397		
Total	340	1020.4			

Table 6. Basic parameter table.

	Estimate	Standard error	t-Statistic	p-value
♯1	0.245794	0.0552653	4.44753	0.0000121198
♯2	−0.473285	0.0566816	−8.34989	2.30183×10^{-15}
♯3	−0.00730259	0.0626806	−0.116505	0.907368
♯4	−0.384589	0.0599047	−6.42001	5.1073×10^{-10}
♯5	−0.244072	0.0613891	−3.97582	0.0000873291
♯6	0.353163	0.0626402	-5.63797	3.87278×10^{-8}
♯7	0.00100867	0.0631682	0.0159681	0.98727
♯8	−0.142987	0.0620628	−2.30391	0.0218891
♯9	0.136853	0.0610078	2.24321	0.0255892
♯10	−0.241157	0.0608958	−3.96016	0.0000929613
♯11	0.167892	0.0599006	2.80285	0.00538436
♯12	−0.128651	0.0605339	-−2.12527	0.0343548
♯13	0.0686504	0.0554803	1.23738	0.216881
♯14	−0.0330885	0.0555248	−0.595923	0.551661
♯15	0.327951	0.0537379	6.10279	3.10786×10^{-9}
♯16	−0.0368829	0.0559783	−0.658877	0.510464
♯17	0.171794	0.0557006	3.08425	0.00222441
♯18	0.357909	0.0566593	6.31687	9.25464×10^{-10}
♯19	0.194009	0.0603826	3.213	0.00145176
♯20	0.0956424	0.0602179	1.58827	0.113244
♯21	0.247097	0.0593144	4.16589	0.0000402579
♯22	0.183056	0.0604346	3.029	0.00265998
♯23	0.335842	0.060323	5.5674	5.60543×10^{-8}
♯24	−0.00644882	0.0597654	−0.107902	0.914143
♯25	0.162157	0.0582993	2.78146	0.00574279
♯26	−0.102138	0.0573372	−1.78136	0.0758319
♯27	0.301513	0.0564203	5.34406	1.76445×10^{-7}
♯28	0.0609255	0.0545446	1.11699	0.264865
♯29	0.0523586	0.0520694	1.00555	0.315414
♯30	0.0408878	0.039727	1.02922	0.304179

4.2. *Non-linear autoregression using machine learning methods.*

We next assume that the autoregressive model $x_{n+31} = F(x_{n+30}, \ldots, x_n)$ is non-linear. Using the previous data, we utilize machine learning methods to obtain a reconstruction of the trajectories of x. Firstly, we interpolate the data with the random forest method with parameters given in Table 7.
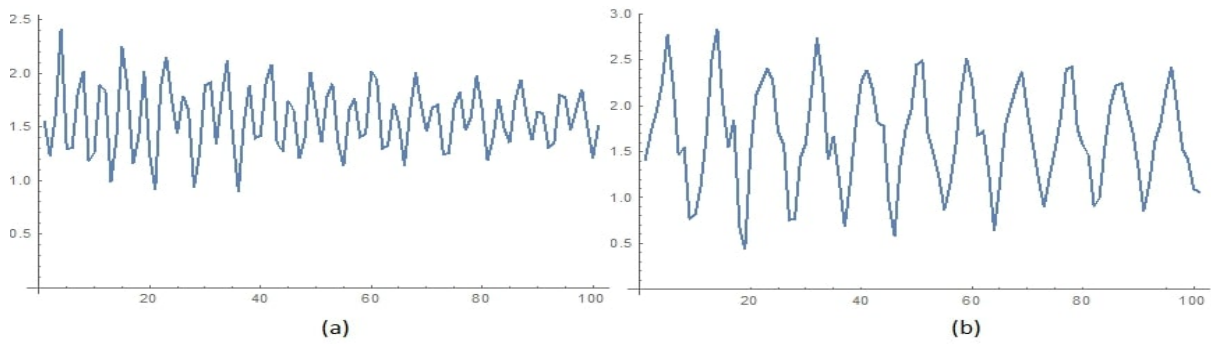
Figure 7. Reconstruction of (a) the first and (b) the second trajectory of *x*.
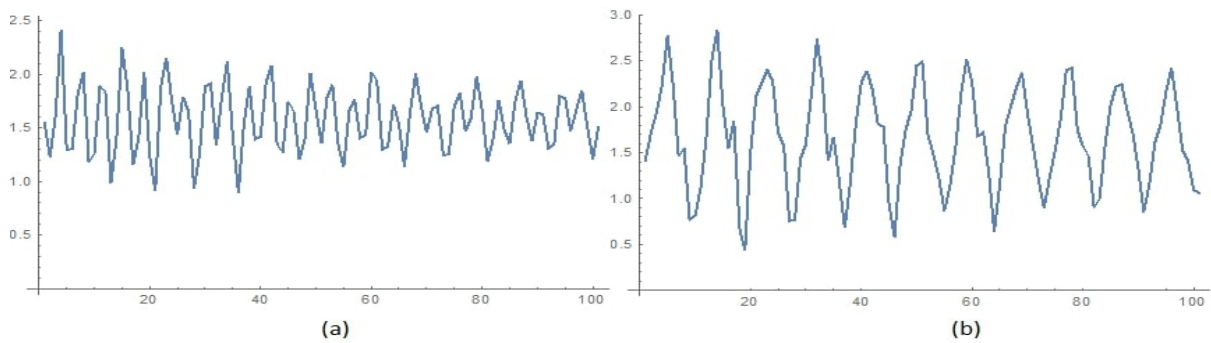
Figure 8. Reconstruction of (a) the first and (b) the second trajectory of *x* using random forest interpolation.

Table 7. Details of the random forest interpolation.

Predictor information	
Method	Random Forest
Number of features	1
Number of training examples	340
Number of trees	50

Figure 8 shows the simulated two trajectories using the random forest predictor.

Table 8. Details of the Gaussian interpolation process.

Predictor information	
Method	Gaussian process
Number of features	1
Number of training examples	340
Assume Deterministic	False
Numerical Covariance Type	Squared Exponential
Nominal Covariance Type	Hamming Distance
Estimation Method	Maximum Posterior
Optimization Method	Find Minimum

Finally, we can interpolate the previous data with the Gaussian process method instead of random forests. The relative information of the used method are described in Table 8. Finally, Figure 9 contains the approximations of the two trajectories of x that can be obtained using this method.
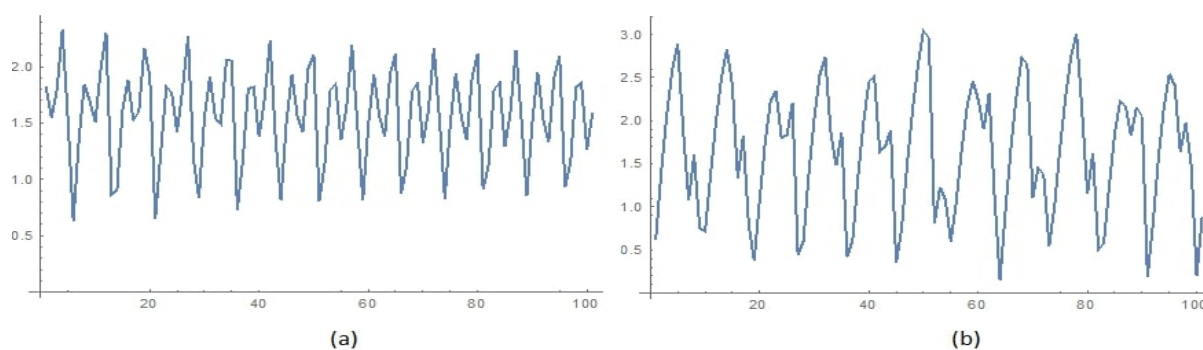
Figure 9. Reconstruction of (a) the first and (b) the second trajectory of x using Gaussian interpolation.

5. Conclusions

Given a chaotic system with a strange attractor, Takens' theorem informs us that we can obtain a structure topologically equivalent to the attractor by means of a delay embedding. Furthermore, it provides us with an upper bound for the required dimension of the embedding. However, this is mainly a theoretical result which may not be very useful in practice, since we may not know the quantities that appear in it. Furthermore, in some cases the estimation of the dimension may be too high.

In this paper, we combine the result of Takens' theorem with ideas from Koopman operators and Extended Dynamic Mode Decomposition. Since EDMD depends on data, it provides the necessary tools for numerical calculations, which are absent from the Takens' framework. In this way, we obtain a numerical method for reconstruction of the trajectories of a chaotic dynamical system. This reconstruction may also lead to simplified versions of the chaotic dynamical system as well as to a better understanding of its dynamics.

In particular, we applied the methodology to a chaotic dynamical system that appears in financial studies. According to Takens' theorem, we consider one observable of the system. This observable may

be one of the state variables. We used the variable x of the system. Similar work can also be executed for the other variables as well as for other observables, if they can provide useful information about the system. Then, we take time-delayed measurements of this observable

$$\mathbf{x}(t) = [x(t), x(t - \Delta t), \ldots, x(t - d\Delta t)]^T .$$

Takens' theorem provides information about the number d of measurements. In particular, when $d > 2n$, the embedding is faithful. Therefore, we may seek for a map $F \colon \mathbb{R}^d \to \mathbb{R}$ such that

$$x(t + 1) = F(\mathbf{x}(t)) = F(x(t), x(t - \Delta t), \ldots, x(t - d\Delta t)).$$

In Section 3, we have tested several choices for the function F. The simplest case is to consider a linear dependence of $x(t + 1)$ form the previous measurement. Other choices that we examined use machine learning methods (random forests, Gaussian interpolation) to find the map F and to reconstruct the trajectory of the variable x of the original dynamical system.

What is the advantage in our approach is the combination of the Koopman-EDMD theory, Takens' theorem and Machine Learning that allows us to use data measurements in order to obtain the orbits of a system exhibiting very complicated behaviour. The key problem with many facets we are faced with is the reconstruction of the trajectories of dynamical systems in general and of financial dynamical systems in particular (Puu (1989)). A first step towards this direction is to use more observables in the Koopman-EDMD theory consisting of the state variables, the observables used in Takens' theorem and (non) linear functions of them.

Koopan Mode Analysis has been applied in energy economics (Georgescu and Mezić (2015)) and in financial economics. More specifically, the findings of the present paper could have practical application in financial trading, extending further the work of Mann and Kutz (2016) who used the Koopman operator in financial trading, developing dynamic mode decomposition on portfolios of financial data. Furthermore, they could be considered together with Hua et al. (2017) proposed novel methodology for high dimensional time series prediction based on the kernel method extension of data-driven Koopman spectral analysis, by researchers on quantitative finance.

Acknowledgments

The authors would like to thank the reviewers for their helpful remarks and comments.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

Ahmad I, Ouannas A, Shafiq M, Pham V, Baleanu D (2021) Finite-time stabilization of a perturbed chaotic finance model. *J Adv Res* 32: 1–14. https://doi.org/10.1016/j.jare.2021.06.013

Chen H, Yu L, Wang Y, Guo M (2021) Synchronization of a Hyperchaotic Finance System. *Complexity* 2021: 7. https://doi.org/10.1155/2021/6618435

Chian AC (2000) Nonlinear dynamics and chaos in macroeconomics. *Int J Theor Appl Financ* 3: 601–613.

Chian AC-L, Rempel EL, Rogers C (2006) Complex economic dynamics: Chaotic saddle, crisis and intermittency. *Chaos Solitons Fractals* 29: 1194–1218. https://doi.org/10.1016/j.chaos.2005.08.218

Eckmann JP, Ruelle D (1985) Ergodic theory of chaos and strange attractors. *Rev Mod Phys* 57: 617–656. https://doi.org/10.1007/978-0-387-21830-4-17

Evstigneev I, Taksar M (2009) Dynamic interaction models of economic equilibrium. *J Econ Dyn Control* 33: 166–182. https://doi.org/10.1016/j.jedc.2008.04.011

Fanti L, Manfredi P (2007) Chaotic business cylces and fiscal policy: an IS-LM model with distributed tax collection lags. *Chaos Solitons Fractals Elsevier* 32: 736–744.

Gao Q, Ma JH (2009) Chaos and Hopf bifurcation of a finance system. *Nonlinear Dyn* 58: 209–216. https://doi.org/10.1007/s11071-009-9472-5

Georgescu M, Mezić I (2015) Building energy modeling: A systematic approach to zoning and model reduction using Koopman Mode Analysis. *Energy Buildings* 86: 794–802. https://doi.org/10.1016/j.enbuild.2014.10.046

Guckenheimer J, Holmes P (1983) Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer-Verlag, New York.

Guégan D (2009) Chaos in economics and finance. *Annu Rev Control* 33: 89–93. https://doi.org/10.1016/j.arcontrol.2009.01.002

Haas L (1998) Stabilizing chaos in a dynamic macroeconomic model. *J Econ Behav Organ* 33: 313–332. https://doi.org/10.1016/S0167-2681(97)00061-9

Hirsch MW, Smale S (1974) Differential Equations, Dynamical Systems, and Linear Algebra. Academic Press, San Diego.

Holyst J, Hagel T, Haag G, Weidlich W (1996) How to control a chaotic economy? *J Evolut Econ Springer* 6: 31–42. https://doi.org/10.1007/BF01202371

Hua JC, Noorian F, Moss D, Leong PHW, Gunaratne GH (2017) High-dimensional time series prediction using kernel-based Koopman mode regression. *Nonlinear Dyn* 90: 1785–1806. https://doi.org/10.1007/s11071-017-3764-y

Jian JG, Deng XL, Wang JF (2009) Globally Exponentially Attractive Set and Synchronization of a Class of Chaotic Finance System. *Lect Notes Comput Sci* 5551: 253–261, Springer, Berlin. https://doi.org/10.1007/978-3-642-01507-6-30

Lorenz HW (1993) Nonlinear Dynamical Economics and Chaotic Motion. Springer, Berlin. https://doi.org/10.1007/978-3-662-22233-1

Ma JH, Chen YS (2001) Study of the bifurcation topological structure and the global complicated character of a kind of nonlinear finance system. *Appl Math Mech* 22: 1240–1251. https://doi.org/10.1007/BF02437847

Ma R, Wu J, Wu K, et al. (2022) Adaptive fixed-time synchronization of Lorenz systems with application in chaotic finance systems. *Nonlinear Dyn* in press. https://doi.org/10.21203/rs.3.rs-1443857/v1

Mann J, Kutz JN (2016) Dynamic mode decomposition for financial trading strategies. *Quant Financ* 16: 1643–1655. https://doi.org/10.1080/14697688.2016.1170194

Mauroy A, Mezić I, Susuki Y (Editors) (2020) The Koopman Operator in Systems and Control. *Lect Notes Control Inf Sci*, Springer. https://doi.org/10.1007/978-3-030-35713-9

Mezić I, Banaszuk A (2004) Comparison of systems with complex behavior. *Physica D* 197: 101–133. https://doi.org/10.1016/j.physd.2004.06.015

Muldoon MR, MacKay RS, Huke JP, et al. (1993) Topology from time series. *Physica D* 65: 1–16. https://doi.org/10.1016/0167-2789(92)00026-U

Ni H, Dong X, Zheng J, et al. (2021) An Introduction to Machine Learning in Quantitative Finance. World Scientific.

Packard NH, Crutchfield JP, Farmer JD, et al. (1980) Geometry from a time series. *Phys Rev Lett* 45: 712–716. https://doi.org/10.1103/PhysRevLett.45.712

Pan X, Wu J (2022) Stochastic stabilization of the chaotic finance system via adaptive fixed-time control. *Chin Control Conf.*, in press. https://doi.org/10.23919/CCC55666.2022.9902830

Puu T (1989) Nonlinear Economic Dynamics, *Lect Notes Econ Math Syst* 336. Springer-Verlag. https://doi.org/10.1007/978-3-642-97291-1-1

Rigatos GG (2017) State-space approaches for Modelling and Control in Financial Engineering. Systems theory and machine learning methods. *Intell Syst Ref Library* 125, Springer. https://doi.org/10.1007/978-3-319-52866-3

Sauer T, Yorke JA, Casdagli M (1991) Embedology. *J Stat Phys* 65: 579–616. https://doi.org/10.1007/BF01053745

Schaffer WM, Kot M (1986) Differential systems in ecology and epiemiology. *Chaos*, Manchester University Press, Manchester.

Stavroglou SK, Pantelous AA, Stanley HE, et al. (2019) Hidden interactions in financial markets. *Proc Nat Acad Sci USA* 116: 10646–10651. https://doi.org/10.1073/pnas.1819449116

Takens F. (1981) Detecting strange attractors in turbulance. *Dynamical Systems and Turbulance, Springer Lecture Notes in Mathematics*, Springer-Verlag, Berlin. 1981: 366–381. https://doi.org/10.1007/BFb0091924

Wijeratne AW, Yi FQ, Wei JJ (2009) Bifurcation analysis in the diffusive Lotka–Volterra system: an application to market economy. *Chaos Solitons Fractals* 40: 902–911. https://doi.org/10.1016/j.chaos.2007.08.043

Yu H, Cai G, Li Y (2012) Dynamic analysis and control of a new hyperchaotic finance system. *Nonlinear Dyn* 67: 2171–2182. https://doi.org/10.1007/s11071-011-0137-9

Zhao XS, Li ZB, Li S (2011) Synchronization of a chaotic finance system. *Appl Math Comput* 217: 6031–6039. https://doi.org/10.1016/j.amc.2010.07.017

AIMS Press