Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 2055655, 11 pages
https://doi.org/10.1155/2022/2055655

Research Article

Identification of Chaos in Financial Time Series to Forecast Nonperforming Loan

Farid Ahmadi,[1] **Mohammad Pourmahmood Aghababa**,[2] **and Hashem Kalbkhani**[2]

[1]*Information Technology Engineering Department, Urmia University of Technology, Urmia, Iran*
[2]*Faculty of Electrical Engineering, Urmia University of Technology, Urmia, Iran*

Correspondence should be addressed to Farid Ahmadi; f.ahmadi@uut.ac.ir

Received 26 October 2021; Revised 24 December 2021; Accepted 19 January 2022; Published 4 February 2022

Academic Editor: Yong Aaron Tan

This paper discusses the importance of modeling financial time series as a chaotic dynamic rather than a stochastic system. The dynamical properties of a financial time series of an economic institution in Iran were analyzed to identify the potential occurrence of the low-dimensional deterministic chaos. This paper applies several classic nonlinear techniques for detecting the chaotic nature of the time series of loan payment portion and proposes a modified nonlinear predictor scheme for forecasting the future levels of the nonperforming loan. The auto mutual information was implemented to estimate the delay time dimension, and Cao's approach, along with correlation dimension methodology, quantified the embedding dimension of the time series. The results reveal a low embedding dimension implying the chaotic nature exists in the financial data. The maximum Lyapunov exponent measure is also adopted to investigate the divergence or convergence of the trajectories. Since positive Lyapunov exponents are revealed, the long-term unpredictability of the time series is proved. Lastly, a modified nonlinear local approximator is developed to forecast the short-term history of the time series. Numerical simulations are provided to illustrate the adopted nonlinear techniques. The results reported in this paper could have implications for commercial bank managers who could use the nonlinear models for early detection of the possible nonperforming loans before they become uncontrollable.

1. Introduction

Banks and financial institutions have played a significant role in balancing the economic life of the people in recent decades owing to the development of the countries and the development of new financial opportunities for the merchants. Lending is one of the primary and popular approaches in such financial organizations aiming to make a loan to somebody on the condition that the amount borrowed is to be returned, usually with an interest fee. However, in some cases, the nonperforming loan (NPL) problem occurs when the borrowed money is not returned in the scheduled period. High levels of NPL mean reducing the income of the banks, which in turn leads to severe economic losses. Therefore, governments have paid more attention to this issue in recent years. Accordingly, a weakening in bank loan services may cause a delay in economic growth and can be a good reason for the economic crisis. It has been argued that NPLs may create economic stagnation and, therefore, can deter economic growth and weaken financial efficacy [1]. On the other hand, high levels of NPLs harm the growth rate of gross domestic product [2]. Given the importance of this issue, the governments and financial sectors should encourage researchers to conduct studies on this issue and banks to implement research results in practice.

As a result, NPL ratios should stand at low or manageable levels before the crisis. The NPL ratio prediction method is necessary to achieve this goal based on previous information. Various studies have attempted to study the relationship between various economic factors and the NPLs to forecast the NPLs. In [1], a heuristic hybrid classification method has been used to predict banks' nonperforming loans using some macroeconomic and bank-specific features. Tang et al. [3] have used a stepwise discrimination algorithm to find essential factors for building distance discrimination and Bayesian discrimination models to determine whether an NPL has a zero or positive recovery rate. In [4], three two-phase mixture models of logistic regression and artificial neural networks have been developed to create an economic distress warning system appropriate for Taiwan's banking business. The application of a neural network predictor in forecasting loan recovery in the Nigerian financial institutions has been reported in [5]. In [6], the principal component analysis (PCA) technique has been adopted for feature selection for prediction models of Chinese bank loans. Then, the models have been assessed using the TOPSIS multicriteria decision-making approach. Saha et al. [7] have utilized machine learning strategies such as ontology, text, data mining, and multiagent approaches to develop a knowledge-determined automatic acquiescence auditing method for bank loans. Some studies have used other techniques, such as nonlinear regression models for loan prediction [8, 9].

Most of the studies mentioned above have tried to check whether a particular financial or banking attribute influences NPLs quotient or not. In most cases, just some measurable quantities which depend on the underlying and usually unknown dynamics of the NPL rate are available. Since many factors affect the ratio of NPLs, the inherent nature of NPLs becomes more complicated and nondeterministic. In such situations, linear estimators that minimize the variance fail to reach the best fitting or forecasting purposes [10]. Moreover, the designer should correctly set many model parameters in the previous nonlinear models (such as neural networks). Therefore, it is hard to predict and interpret the long-term future of NPLs with limited features. The complex behavior of the NPLs cannot be easily modeled using the common linear or nonlinear statistical approaches, such as autoregressive methods and neural networks. Thus, it would be better to use alternative nonlinear powerful techniques that utilize the inherent attributes of the NPLs' nature, such as chaos [11]. In the past few decades, the concept of nonlinear dynamical systems and chaos theory has changed the treated manner of financial systems. The complicated behavior of NPLs can be interpreted using a chaotic system (as well as chaos theory) with high sensitivity to the initial conditions.

Chaos is a particular version of nonlinear dynamics that possesses unique attributes such as incredible sensitivity to the initial states (a tiny change in the starting point will produce a significant diversity in the future), broad Fourier transform spectrum, irregular attractors (the states are locally unbounded but globally bounded), and fractality property. A chaotic phenomenon is neither stochastic nor random; it is a deterministic system in which some equations are available to determine its behavior. Moreover, although the chaotic time series are not periodic and random, they exhibit a sense of order and pattern. Such unique and complicated appearances can be detected using nonlinear techniques, including phase space reconstruction, false nearest neighbor (FNN) algorithm, correlation dimension method, and Lyapunov exponent. Consequently, many real-world applied systems that may seem to have random nature can possess a nonlinear deterministic and potentially chaotic behavior. In this line, with the introduction of the ability to model and predict the future of chaotic time series through nonlinear deterministic system theory, many researchers have attempted to study and detect chaos in time series of multidisciplinary fields [12, 13]. Also, complex nonlinear dynamical modeling and analysis have recently become more important in advanced science approaches [14–16] and [17]. However, a chaotic time series analysis is still a good choice in many economic-related issues.

In [18], Novikov's theorem has been utilized to model the complicated dynamics of noisy credit risk contagion with time delay, and the Hopf bifurcation and chaotic behaviors are evaluated. In [19], some numerical approaches have been applied to discover Hopf bifurcation, inverse bifurcation, and chaos phenomena in the credit risk contagion dynamics. Lahmiri [20] has investigated the fractal inherence and chaotic behavior in returns and volatilities of family business companies of Morocco, using Hurst exponent and an autoregressive model. In [21], the phase synchronization method has been introduced for analyzing the chaotic behavior of stock price and index movements in crisis stages. For identifying the quality of similarity measure of financial time series, three techniques including information categorization approach, reconstructed phase space clustering strategy, and system methodology with squared Euclidean distances have been used in [22]. In addition to these works, some research in the literature aims to predict the chaotic financial time series. In [23], a self-organizing map neural network along a recommender system has been proposed to cluster and predict stock price time series. Yang and Lin [14] have applied empirical mode decomposition and phase space reconstruction methods combined with extreme learning machines for predicting financial exchange rates' time series forecasting. The artificial neural network method is the main and mostly applied technique in the literature for predicting the future of the financial chaotic time series [24–27]. To the best of our knowledge, no single review article detecting and predicting chaos in NLPs' time series has been published; therefore, it will be addressed in the present work.

This research aims to investigate the financial time series behavior of the primary interest-free institution of Iran named Omid Entrepreneurship Fund (OEF). The main purpose is to identify the potential occurrence of low-dimensional deterministic chaos in the financial time series of OEF and to propose an efficient forecasting technique to predict the future events of NPLs. The dataset is collected during the last five years of the OEF activity. We adopt the loan payment percent (LPP) as the time series to be analyzed. We try to predict the future of the LPP with a modified chaos-based nonlinear methodology. As a result, the predicted values can be utilized to detect and predict NPLs (an NPL can be determined by an LPP less than a specific value in a given time) before they become uncontrollable. The superposition of all exogenous and endogenous variables affects the loan payment percent. Hence, in this paper, the effects of exogenous and endogenous were considered indirectly. The delay time dimension is discovered using auto mutual information to investigate the chaos in the time series. It is employed to rebuild the irregular attractors. The embedding dimension is calculated via Cao's technique and correlation dimension approach. The quantity of the sensitivity to the initial states is computed with the help of the largest Lyapunov exponent (LLE) principle. The given positive Lyapunov exponents assure the exponential deviation of the trajectories and, therefore, the unpredictability of the time series. Finally, a modified nonlinear predictor is realized to forecast the financial time series. All the mentioned approaches are illustrated using computer simulations.

The rest of this article is organized as follows. In Section 2, the nonlinear chaos-based techniques are presented. Section 3 deals with the description of a modified prediction scheme. In Section 4, some computer simulations are carried out. Finally, concluding remarks are provided in Section 5.

2. Nonlinear Techniques for Analyzing Chaotic Time Series

This section discusses the adopted nonlinear techniques for chaos detection in the financial time series.

2.1. Phase Space Rebuilding. Since the chaotic dynamics of an irregular time series are not precisely known, a phase space estimation of the state space of a time series can be effectively used to rebuild the equivalent irregular attractor. To this end, the time series is converted into the geometry of a single moving point along a chaotic trajectory, where each of its points corresponds to a state of the analogous chaotic system. As a result, the phase space rebuilding can be inferred as a multidimensional depiction of a single-dimensional nonlinear time series. A delay-based applied method for rebuilding the phase space has been introduced by Takens [28]. The main idea of the Takens method is that the matching trajectories of the systems with a set of naturally deterministic dynamics can reach toward the subset of the phase space (i.e., the attractor). So, that technique assures that an m-dimensional space can be created to embed the original behavior of a given time series $X_i, i = 1, 2, \ldots, N$, in which each element of the state vector is obtained via the delay coordinates as below [28]:

$$Y_j = \left(X_j, X_{j+\tau}, X_{j+2\tau}, \ldots, X_{j+(m-1)\tau}\right), \tag{1}$$

where $j = 1, 2, \ldots, N - (m-1)\tau/\Delta t$, m is named the embedding dimension ($m \geq d$ where d shows the dimension of the attractor), τ denotes the delay time, and Δt represents the sampling time. Computational errors caused by a finite precision arithmetic allow to consider one pseudotrajectory computed for a sufficiently large time interval [29–31].

2.2. Time Delay Estimation. Two main approaches can carry out a proper estimation of the time delay τ. In the first strategy, one can calculate the autocorrelation function of the time series and choose the first zero-crossing time. In this strategy, the $X_{i+\tau}$ sample can be fully decorrelated from the X_i sample, once the autocorrelation function reaches zero at spots the past point [32, 33]. This technique reproduces only linear features of the dynamics and usually requires supplementary data. At the same time, the other methodology involves a nonlinear autocorrelation function named mutual information (MI) to compute the delay from the time series [34]. Shannon's information theory inspires the main motivation of this technique to produce the information accomplished from examinations of one random event on another using the MI criterion. The MI is a nonlinear equivalent to the correlation function, and both linear and nonlinear reliance among two time series can be measured by it. Once the MI is adopted for time-delayed translations of the identical sequence, it is named auto MI (AMI).

Usually, the MI recognizes the quantity of information a signal gives regarding the other signal. Thus, the AMI calculates the approximate degree of forecasting $X_{i+\tau}$ from X_i. From an information-theoretic point of view, the AMI discovers how the measurements X_i are joined $X_{i+\tau}$. The following formula gives the AMI [34]:

$$I(\tau) = \sum_{X_i, X_{i+\tau}} P(X_i, X_{i+\tau}) \log_2\left(\frac{P(X_i, X_{i+\tau})}{P(X_i)P(X_{i+\tau})}\right), \tag{2}$$

where i denotes total sample number, $P(X_i)$ and $P(X_{i+\tau})$ show the marginal probabilities for measurements X_i and $X_{i+\tau}$, respectively, and $P(X_i, X_{i+\tau})$ gives their connection probability density for measurements X_i and $X_{i+\tau}$.

Remark 1. A delay time τ that minimizes $I(\tau)$ for $t = \tau$ and $X_{i+\tau}$ appends the highest information on X_i has the optimal value.

2.3. Embedding Dimension. The minimum number of the state variables needed to display the system behavior is called the embedding dimension m. The Grassberger–Procaccia (GP) [35], the singular value decomposition (SVD) [36], the FNN [37], and Cao's scheme [38] are the main approaches for obtaining the minimum embedding dimension from a scalar time series. The delay coordinates at a specified time delay τ are used in the GP method to rebuild the dynamics of a scalar time series in an embedding space of dimension m. Although this procedure is data demanding and subjective and consumes more time in the simulation, it can determine the time series's chaotic and/or random nature. With the help of singular values of embedding, the quantity of variance of the trajectory's projection on the orthogonal directions in the embedding space is used in the SVD method. The minimum dimension is computed by the number of directions the rebuilt trajectory sees and determined by the large singular values. The main drawback of the SVD scheme is its subjectivity to the number of singular values.

The FNN algorithm is inspired by the fact that the orbits of a chaotic attractor cannot cross or go beyond each other. In contrast, when a lower dimension than the adequate characterizes an irregular attractor, a junction and/or an overlap occurs. In other words, the FNN algorithm relies on that if a too low embedding dimension is chosen, the points distant from each other in the original phase space will be closed in the rebuilt phase space. Subjectivity and the need for supplementary data are weaknesses of this approach.

The drawbacks of the algorithms above are tackled using Cao's strategy. This scheme adopts a scalar time series $X_1, X_2, \ldots, X_N$ and the delay time vector (1) with the following norm [38]:

$$a(i,m) = \frac{\left\| y_i(m+1) - y_{n(i,m)}(m+1) \right\|}{\left\| y_i(m) - y_{n(i,m)}(m) \right\|}, \quad i = 1, 2, \ldots, N - m\tau, \tag{3}$$

where $y_i(m+1)$ denotes the ith reconstructed vector with embedding dimension $m+1$, i.e., $y_i(m+1) = (X_i, X_{i+\tau}, \ldots, X_{i+m\tau})$, $1 \le n(i,m) \le n - m\tau$ stands for an integer such that $y_{n(i,m)}(m)$ is the nearest neighbor of $y_i(m)$ in the m-dimensional rebuilt phase space, and $\| y_k(n) - y_l(n) \| = \max_{0 \le j \le n-1} |x_{k+j\tau} - x_{l+j\tau}|$.

Remark 2. If $y_{n(i,m)}(m)$ is equal to $y_i(m)$, the second nearest neighbor is adopted instead of it.

In Cao's and FNN approaches, any two points represent true neighbors that their nearness in the m-dimensional rebuilt phase space guarantees that they are still near in the $(m+1)$-dimensional rebuilt phase space. If this condition is not met, the two points will be interpreted as false neighbors. In a perfect embedding, no false neighbors should be supposed. In [32], a false neighbor has been recognized by checking whether $a(i,m)$ is larger than some threshold rate. Nonetheless, it is clear from (3) that this requires the derivative of the original signal. Hence, different threshold values are essential for different phase points i implying that it is difficult to gain an appropriate threshold value independent of the values of m, X_i, and the time series.

To circumvent the issue above, Cao has applied the average value of all $a(i,m)$ s as another quantity as follows:

$$E(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} a(i,m). \tag{4}$$

The variation of $E(m)$ from m to $m+1$ can be evaluated using the following formula [35]:

$$E_1(m) = \frac{E(m+1)}{E(m)}. \tag{5}$$

Once the time series appears from a chaotic attractor, $E_1(m)$ ends variation when m is greater than a fixed value m_0. In this situation, $m = m_0 + 1$ is taken as the minimum embedding dimension.

2.4. Largest Lyapunov Exponent Measure. It is well known that once a time series is susceptible to the initial states, it will be unpredictable, at least for the long term. The highly sensitive time series have divergent exponential trajectories varying with small fluctuations of the initial states. A mean norm of this divergence and the unpredictability of a chaotic time series are given by the Lyapunov exponent measure, in which it expresses the rate of division of infinitesimally close states. The LLE calculates the deviation of close trajectories in the phase space. Therefore, a positive Lyapunov exponent implies the occurrence of chaos.

Suppose that s_{t1} and s_{t2} are two points in two trajectories in the state space with $\| s_{t1} - s_{t2} \| = \partial_0 \ll 1$. If Δt time steps are continued, one has $\partial_{\Delta t} \cong \| s_{t1+\Delta t} - s_{t2+\Delta t} \|$, $\partial_{\Delta t} \ll 1$ and $\Delta t \gg 1$, which means that the initial division ∂_0 will diverge the trajectories with an exponential rate, $\partial_{\Delta t} \cong e^{\lambda \Delta t} \partial_0$, where λ is the Lyapunov exponent [39]. As a result, a positive λ assigns an exponential divergence of the closed trajectories, implying a chaotic behavior. On the other side, dissipative and nonconservative time series show negative Lyapunov exponents, and their trajectories will attain a stable equilibrium point or a periodic orbit. Also, the Lyapunov exponent of the conservative time series is equal to zero.

In work [40], a numerical algorithm has been given for determining the LLE of the scalar time series. The algorithm looks for every neighbor within a neighborhood of command trajectory and gets the average distance of neighbors and the command trajectory as a function of time. This algorithm computes a stretching factor $S(\tau)$ whose slope is equal to the LLE. The methodology in [39] takes a t from the set $T_\tau = \{m, m+1, \ldots, T - \tau\}$ to find $U_t = \varepsilon-$ neighborhood of an arbitrary point X_t in the time series. Afterward, for all $i \in U_t$, the distance of X_t and a neighbor of it is computed as $|X_{t+\tau} - X_{i+\tau}|$, and the logarithm of the mean of these distances is found. This process is repeated for all $t \in T_\tau$, and the stretching factor is achieved.

$$S(\tau) = \operatorname*{mean}_{t \in T_\tau} \left(\ln \left(\operatorname*{mean}_{i \in U_\tau} \left(\left| X_{t+\tau} - X_{i+\tau} \right| \right) \right) \right). \tag{6}$$

The other algorithm for estimating the LLE of a scalar time series has been introduced by Rosenstein et al. [41]. This algorithm utilizes all the information of the entire dataset rather than relying on one trajectory. Once the phase space is built by supposed τ and m, a point X_{n0} is taken, all neighbor points X_n closer than a distance r are recognized, and the average distance from that point between them is calculated. This process is reiterated for N points along the trajectory to find an average value of S as follows:

$$S = \frac{1}{N} \sum_{n_0=1}^{N} \ln \left(\frac{1}{\left| U_{X_{n_0}} \right|} \sum \left| X_{n_0} - X_n \right| \right), \tag{7}$$

where $|U_{X_{n_0}}|$ denotes the number of neighbors recognized around the point X_{n_0}. The diagram of the stretching factor against time $t = N\Delta t$ appears an arc with a linear increase at the start, continued by an approximately smooth area.

2.5. Surrogate Data. One of powerful tools for the nonlinearity examination is surrogate data test [12, 42], in which the null hypothesis indicates that the observed spatial series is produced by a Gaussian (linear) process with a possibly nonlinear static transform. This method considers the mean, the standard deviation, the cumulative distribution function, and the power spectrum of the original data. The surrogate approach produces substitute data with the identical probabilistic organization as the original data [12].

The iterative amplitude-adjusted Fourier transform (IAAFT) has been introduced in [42] to produce the surrogate datasets. In this approach, the created surrogates rely on the idea of generating constrained realizations, where the measurable features of the spatial series are taken into account rather than the basic model equations [12]. The null hypothesis of underlying Gaussian linear stochastic processes can also be formulated, denoting that randomized samples can be created by generating orders with the same linear attributes as the observed data but which are otherwise random [12]. In the surrogate data test, if the squared amplitudes of the discrete Fourier transform are adopted to denote the linear characteristics of the data, the surrogate spatial series will be produced by multiplying the Fourier transform of the data by random phases and subsequently transforming back to the time domain [12]. An appropriate subsequence of the data is chosen before creating the surrogates to avoid the periodicity artifact from leading to false nonlinearity test outcomes [12]. Moreover, windowing and zero-padding techniques can be applied to suppress the edge effects' problems. However, these methods degrade the invertibility of transform and, therefore, cannot be effectively applied for the phase randomization of the surrogates [12].

In the surrogate data test, the IAAFT is applied to preserve the probability density function and the correlation structure (and therefore the power spectrum) of the original data by iteratively minimizing the deviation. The main procedure of the adopted algorithm is given below [12, 43].

(a) A sorted record of the original spatial series $\{s_n\}$ and the squared amplitudes of its Fourier transform, $S_k^2 = |\sum_{n=0}^{N-1} e^{j2\pi \mathrm{kn}}|^2$, are saved.

(b) The data are randomly shuffled (without substitution) $\{S_n^{(0)}\}$ to corrupt any nonlinear associations and correlations.

(c) The Fourier transform of $\{S_n^{(i)}\}$ is calculated, and its squared amplitude is substituted by S_k^2. The result is transformed back to the time domain.

(d) The resulting series is ranked in order, and each value is replaced with the original series value with an equal rank. This work corrects the probability density function of the data and modifies the power spectrum once more.

(e) Steps (c) and (d) are repeated to achieve a given accuracy.

3. Nonlinear Predictor

In nonlinear dynamics theory, although chaos is deterministic, it cannot be cast further than short intervals. The average forecast horizon of a chaotic time series can be reached by the LLE norm as follows [44]:

$$\Delta t_{\max} = \frac{1}{\lambda_{\max}}. \tag{8}$$

In recent years, some approximate linear and/or nonlinear algorithms have been presented in the literature to forecast the future of the chaotic time series. In [45], a simple nonlinear technique has been provided, restated as follows. First, the delay time of the time series $X_i, \quad i = 1, 2, \ldots, N$, is calculated to rebuild the phase space. Then, the irregular attractor is established with an embedding dimension m, and the following m-dimensional map f_T is applied to model the dynamics of the time series.

$$Y_{j+T} = f_T\left(Y_j\right), \tag{9}$$

where Y_j and Y_{j+T} are vectors of dimension m in which they denote the state at the current time j and the state at the future time $j+T$, respectively.

The next step is to get the help of the observed time series to discover a proper estimation of f_T. In the local approximation procedure [44], a locally piecewise structure is adopted to build the dynamics in the embedding space. Moreover, the domain is divided into some local neighborhoods, and the model is accomplished for each neighborhood separately, resulting in a different f_T for each subset. Hence, the dynamics of the system are governed part by part, and the complexity of f_T is considerably reduced without affecting the accuracy of the forecast.

The deviation of the trajectory concerning the time should be estimated to forecast in the m-dimensional space. Regarding the relation of the X_t and X_{t+p} points, the future of the system at the time p on the irregular attractor can be estimated via a nonlinear function F as follows:

$$X_{t+p} \cong F\left(X_t\right). \tag{10}$$

The primary assumption of this forecast algorithm is that the variation of X_t with time on the irregular attractor is identical to those of close points $(X_{T_h}, \quad h = 1, 2, \ldots, n)$. Afterward, X_{t+p} is built by the order d polynomial $F(X_t)$ as follows.

$$\begin{aligned} X_{t+p} \cong{} & f_0 + \sum_{k_1=0}^{m-1} f_{1k_1} X_{t-k_1\tau} + \sum_{\substack{k_2=k_1 \\ k_1=0}}^{m-1} f_{2k_1k_2} X_{t-k_1\tau} X_{t-k_2\tau} \\ & + \cdots + \sum_{\substack{k_d=k_{d-1} \\ \cdots \\ k_1=0}}^{m-1} f_{\mathrm{d}k_1k_2\ldots k_d} X_{t-k_1\tau} X_{t-k_2\tau} \cdots X_{t-k_d\tau}. \end{aligned} \tag{11}$$

In [44], it has been suggested that using n of X_{T_h} and X_{T_h+p} with known values, the coefficients f can be reached by the following equation:

$$X \cong \mathrm{Af}, \tag{12}$$

where $X = (X_{T_1+p}, X_{T_2+p}, \ldots, X_{T_n+p}), \quad f = (f_0, f_{10}, f_{11}, \ldots, f_{1(m-1)}, \ldots, f_{d(m-1)(m-1)\cdots(m-1)})$ and A is a $n(m+d)!/m!d!$ Jacobian matrix described below.

$$A = \begin{bmatrix} X_{T_1} & X_{T_1-\tau} & \cdots & X_{T_1-(m-1)\tau} & X_{T_1}^2 & \cdots & X_{T_1-(m-1)\tau}^2 \\ X_{T_2} & X_{T_2-\tau} & \cdots & X_{T_2-(m-1)\tau} & X_{T_2}^2 & \cdots & X_{T_2-(m-1)\tau}^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ X_{T_n} & X_{T_n-\tau} & \cdots & X_{T_n-(m-1)\tau} & X_{T_n}^2 & \cdots & X_{T_n-(m-1)\tau}^2 \end{bmatrix}. \tag{13}$$

However, in this paper, the idea of memory usage is proposed. In our approach, the Jacobian matrix A not only is constructed by m points of the current attractor, but also benefits the other m points of the past attractors to increase the accuracy of the approximate model. Thus, the size of the Jacobian matrix A is modified as $n(K(m+d))!/(Km)!(Kd)!$, where $1 \leq K \leq m$ is selected as an integer. The rows matrix A must approve the following inequality to attain a sure solution:

$$n \geq \frac{(K(m+d))!}{(Km)!(Kd)!}. \tag{14}$$

4. Data and Numerical Results

In this section, a brief description of the adopted data is given. Then, comprehensive numerical simulation and analysis are provided to verify the data's complex behavior and forecast the short-term future of the chaotic financial time series.

4.1. Financial Data. Recently, some microfinance institutes have the mission of providing limited loans with no (or at least a minimum) interest. The target population of such interest-free organizations is low-income people who lack access to the financial services of other banks or traditional financial institutions. The essential condition for requesting an interest-free loan is that the recipient must prove setting up a small-scale enterprise. The essential goal of such microfinance institutions is to help low-income people get better access to financial services and finance small or medium projects. Omid Entrepreneurship Fund (OEF) is the biggest and the most important interest-free institution in Iran. OEF has at least one branch in each state of Iran. We have taken the five-year loan payment percent of two branches of OEF that have the highest activity in the considered time horizon and call them B_1 and B_2 branches. For the considered loans, the due of the first installment is in the range of years 2011–2017. The considered period does not contain any crisis years and other shocks. Since an NPL is defined as a loan with no payment for at least 18 months for the OEF policies, the two last years' data (which do not contain an NLP) are removed. The time series of the two time series are depicted in Figure 1. Furthermore, the significant statistical attributes of these time series are given in Table 1.

4.2. Phase Space. We choose $m = 2$ and $\tau = 1$ to rebuild the phase space and make an illustrative attractor. In this case, the irregular attractor will be projected to the plane $\{X_i, X_{i+1}\}$. Figure 2 depicts the phase space reconstruction of the B_1 and B_2 time series. One sees that an irregular attractor occurs in both the phase space plots. Moreover, these plots deny the requirement for a stochastic modeling approach for the time series because the attractors are in well-defined regions, implying that the deterministic chaos can effectively elucidate the system dynamics.

4.3. Time Delay Assignment. The time delay is estimated using the AMI algorithm with a delay time belonging to [1, 30]. The results for the two datasets are illustrated in Figure 3. Finding the first local minima on the diagrams is needed to estimate an appropriate time delay. This happens in the delay time of 4 and 3 for B_1 and B_2, respectively.

4.4. Embedding Dimension Estimation. Cao's algorithm is implemented to find the minimum value of the embedding dimension to determine the sufficient embedding dimension for the phase space reconstruction. The maximum value of the embedding dimension is assigned equal to 50 to achieve a good result. Then, the criterion $E_1(m)$ is depicted versus the variation of m. The simulation results are plotted in Figure 4. Figure 4(a), which stands for B_1 time series, shows that $E_1(m)$ end varies $m_0 = 16$. Thus, based on Cao's technique, the minimum embedding dimension for B_1 is $m = m_0 + 1 = 17$.

On the other hand, as Figure 4(b) shows, the curve $E_1(m)$ ends vary with $m_0 = 15$ in the case of B_2. So, the minimum embedding dimension of these data equals to $m = m_0 + 1 = 16$. The most important issue is that since $E_1(m)$ does not vary after some finite value of m_0, both the time series B_1 and B_2 are originated from an irregular attractor, implying the existence of chaos.

4.5. Largest Lyapunov Exponent. The computation of the LLE is carried out using the algorithm proposed in [41]. The time delay and minimum embedding dimension parameters computed in the previous subsections are used in the numerical simulation. The curves for the stretching factor S versus the number of points N are displayed in Figure 5. One can see that linear increased regions with some fluctuations placed over the linear parts occur. It is noted that owing to the stretching factor as an average value of the local stretching or shrinking rates in the irregular attractor, some fluctuations typically emerge, and the different rates cannot be continually flattened by the averaging process of the algorithm [46].

Accordingly, the slopes of the curves about the linear regions are determined using a least-squares line fitting approach aiming to achieve the LLEs of the time series. The LLEs of the B_1 and B_2 time series are achieved as 0.0205 and 0.0178, respectively. These positive values of the LLEs confirm the exponential divergence rate of the trajectories and, then, the chaotic behavior of the financial time series. Besides, the inverse of the LLEs is computed to get the forecast horizon. In this situation, it is revealed that the forecast horizons are 48 and 56 samples for B_1 and B_2, respectively. The predictions over these values are subject to extreme uncertainties.

4.6. Surrogate Data. The potential nonlinearity of the time series can be tested for the surrogate data procedure. At $\alpha = 1\%$ the significance level, to gather $1/\alpha - 1 = 99$ surrogate time series, the iterative algorithm given in the previous subsections is adopted to verify that if the original data come from a linear

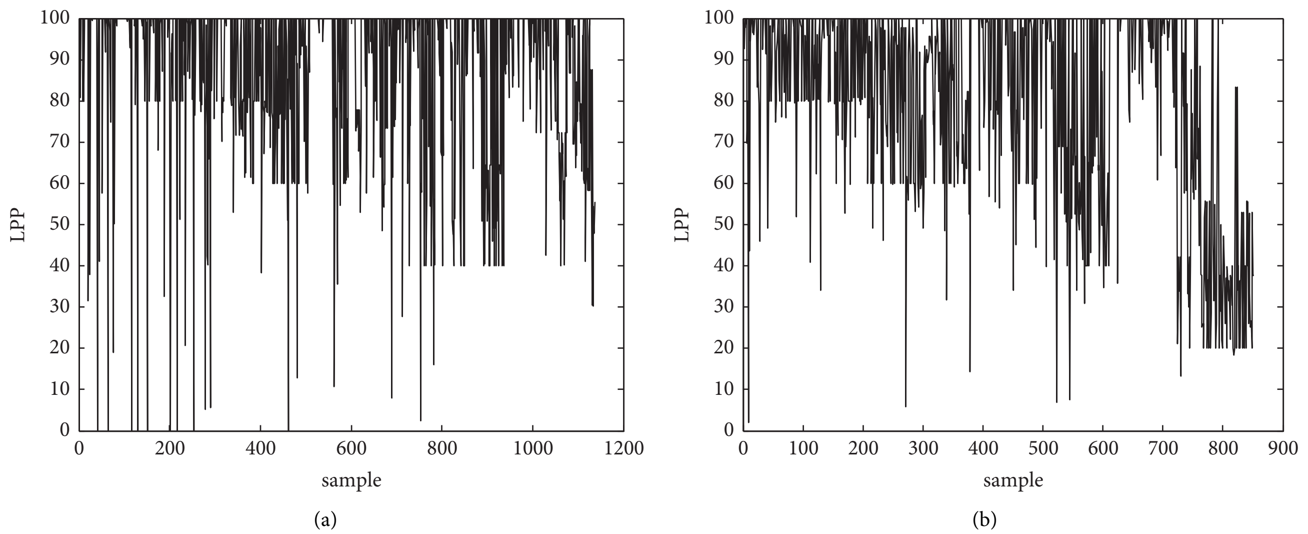


Figure 1: Time series of (a) B_1 and (b) B_2.

Table 1: Statistical attributes of the time series.

Statistical attribute	B_1	B_2
Number of data	1138	851
Mean	87.9281	80.0138
Median	100	92.3100
Min	0	1.9300
Max	100	100
Standard deviation	19.8204	24.4088
Variance	392.8466	595.7916
Skewness	−1.9952	−1.0641
Kurtosis	7.2104	3.0071

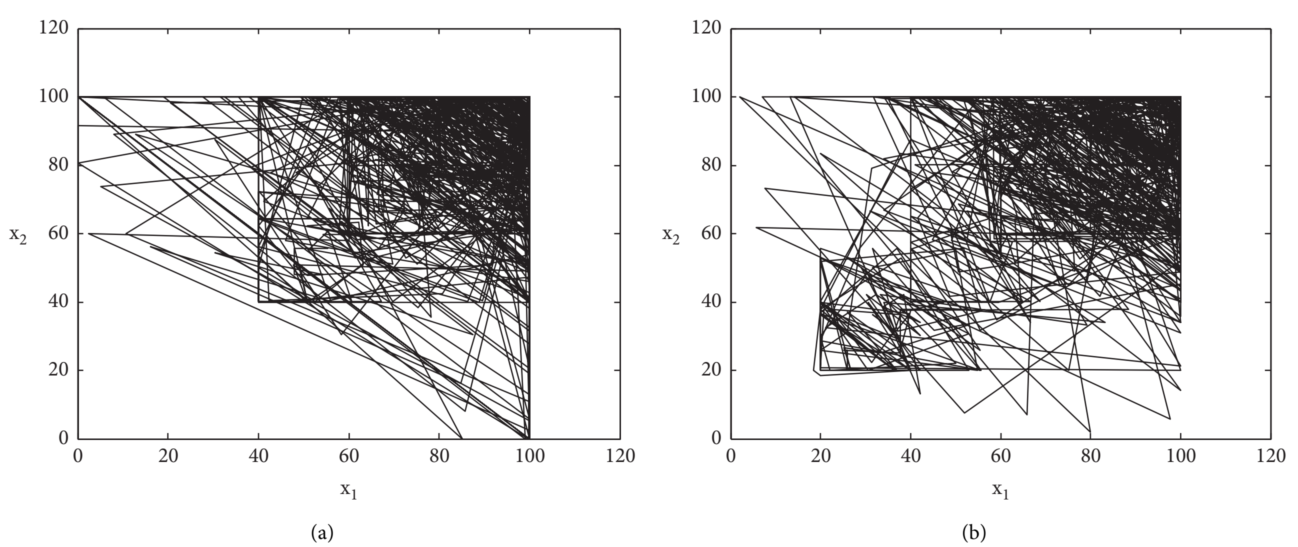
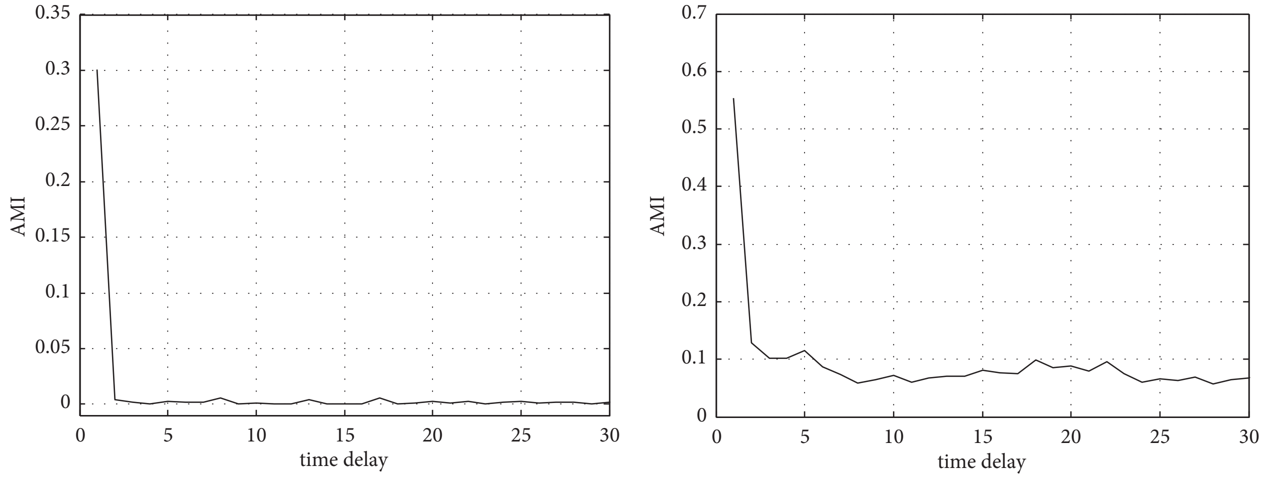


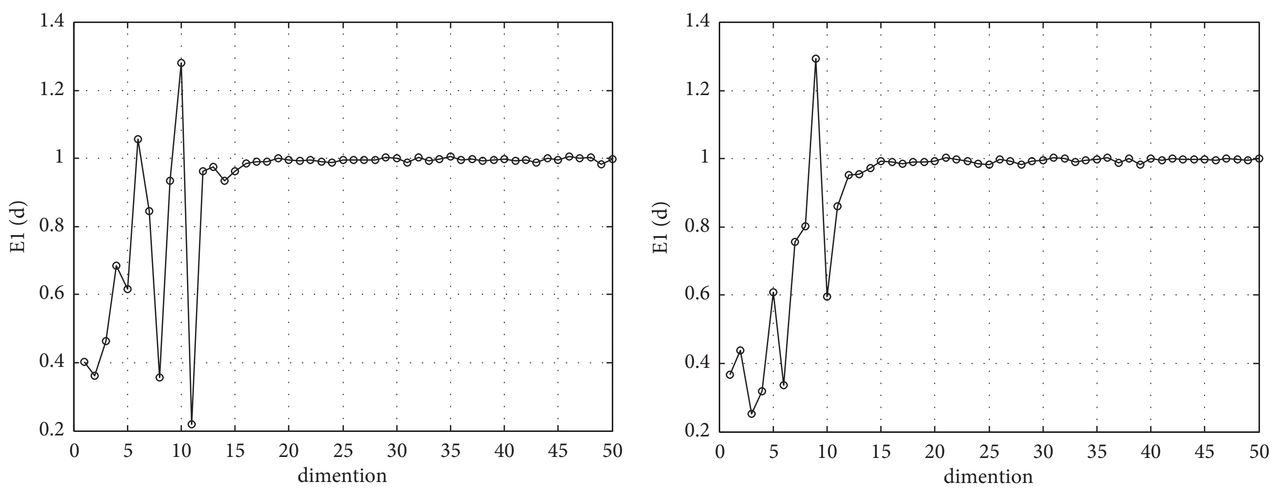
Figure 2: Phase spaces of the time series (a) B_1 and (b) B_2.

stochastic process or not. In this regard, the nonlinear prediction error examination is applied for the test statistic. The diagram of the nonlinear prediction errors of 99 surrogates, and the original data are plotted in Figure 6. The lower thick long line stands for the prediction error, and the other thin lines show the other 99 surrogates, indicating that the null



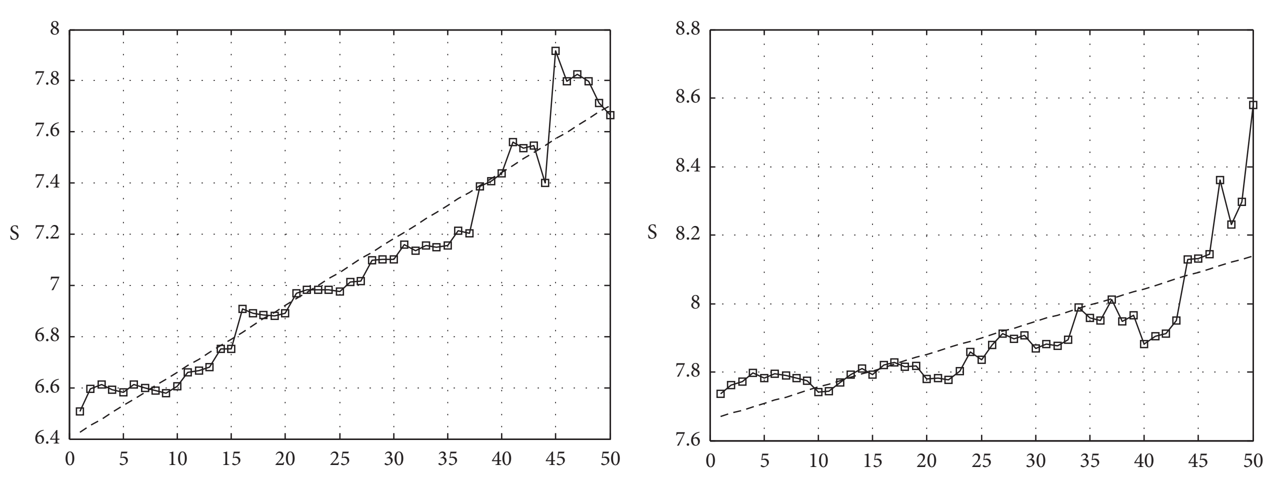
(a) (b)

Figure 3: AMI for the time series (a) B_1 and (b) B_2.



(a) (b)

Figure 4: Cao's results for the time series (a) B_1 and (b) B_2.



(a) (b)

Figure 5: LLEs for the time series (a) B1 and (b) B2.

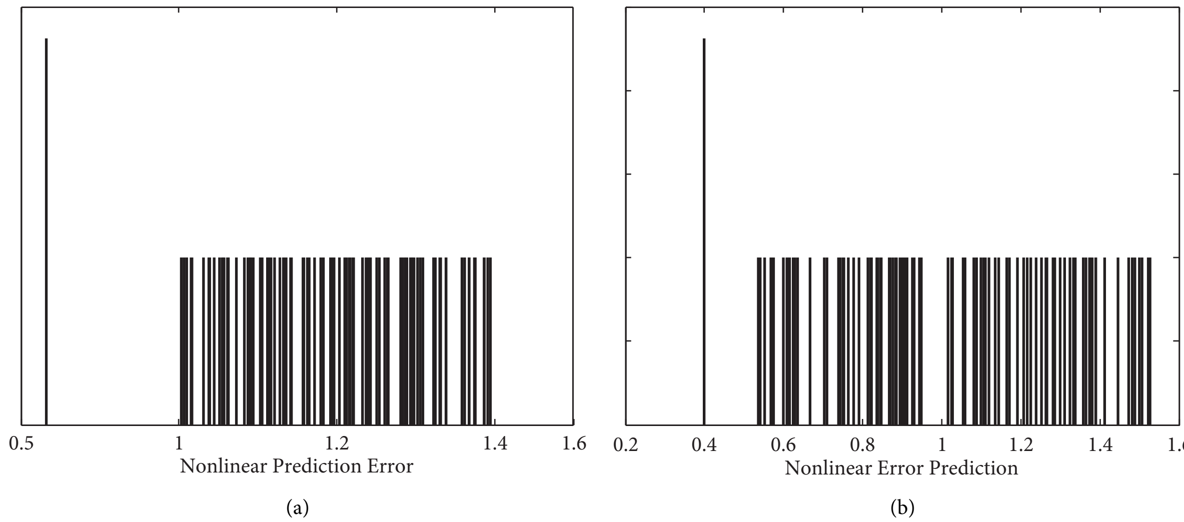


Figure 6: Surrogate data test for the time series (a) B_1 and (b) B_2.

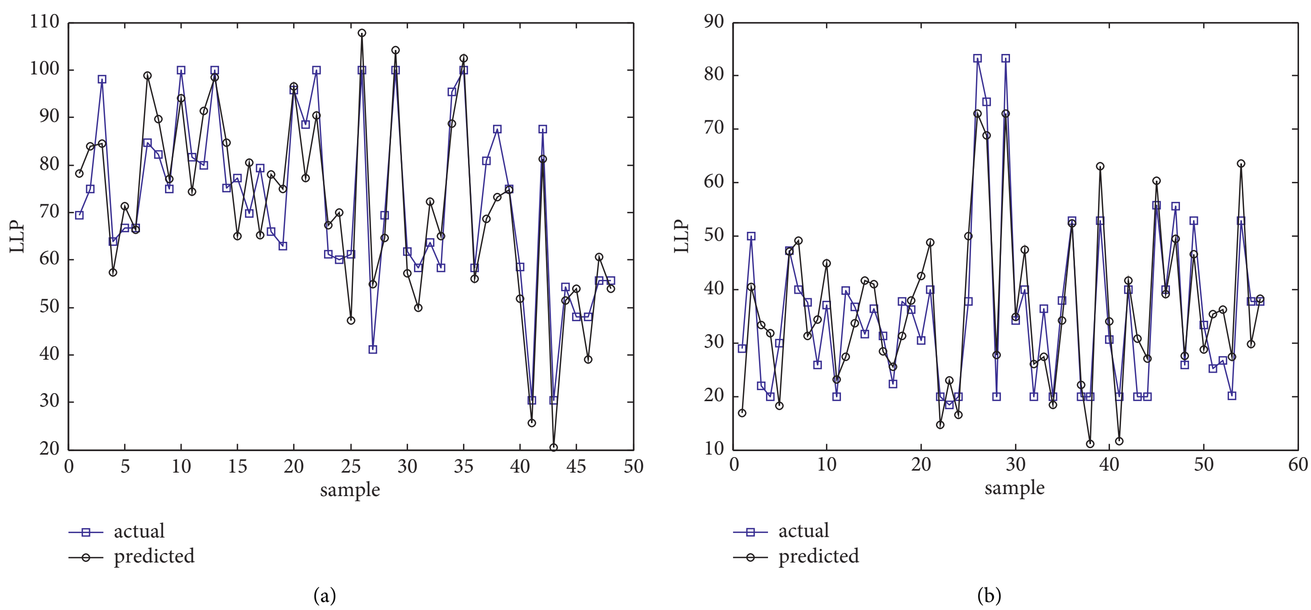


Figure 7: Forecast results for the time series (a) B_1 and (b) B_2.

hypothesis is declined at the 99% significance level. This means that the hypothesis that the data originality is discarded from a linear stochastic process. Therefore, the surrogate data test further validates that the correlation dimension's convergence occurs due to the low-dimensional dynamics that are dominant in the chaotic systems; i.e., it is not due to the linear stochastic nature of the data.

4.7. Prediction. We attempted to forecast the future of the LPP with the modified chaos-based nonlinear prediction method. The performed analysis in the prior sections has proved that the financial time series exhibit a chaotic behavior with a long-term unpredictability feature. On the other hand, the modified local forecasting algorithm can model their future values quickly. The inverse of LLEs suggests that the maximum window for accurate forecasting is 48 and 56 samples for B_1 and B_2, respectively. The motivation for such prediction is clear: the developments of many financial processes have become ever more demanding for the exact short-term prediction. It is noticed that the partial predictability of a chaotic time series owes to its high sensitivity to initial states. This fact indicates that the information is built in the whole time series and that short-time predictions are reasonable. The forecasting results of the time series (with $K = 2$) and the corresponding actual values are depicted in Figure 7. It is viewed that the forecasting results are satisfactory, and they follow the actual data reasonably accurately, except for extreme value. As a result, the predicted values can be utilized to identify and predict NPLs (an NPL is an LPP whose value is less than a particular fixed constant in a specified time) before they become uncontrollable. Therefore, these predictions can help the managers of economic institutions to develop the necessary policies regarding the budget and the remainder of the loans.

5. Concluding Remarks

This paper has introduced the concept of chaotic predictions over short periods for the loan time series. The main idea is that if one can forecast the future behavior of the loan payment percent, then the nonperforming loans will be identified. So, the policies can be modified according to the current state of the budget. Various nonlinear dynamic methods have been realized to identify the existence of low-dimensional chaos in the data to implement this idea. First, the well-known AMI algorithm has discovered the delay time to rebuild the possible irregular attractors. Afterward, the dimensionality of the trajectories was detected using Cao's technique. A low embedding dimension has validated the low-dimensional chaos in the financial data. Accordingly, based on the largest Lyapunov exponent norm, it has been revealed that the time series is susceptible to tiny fluctuations of the initial states. This proves the exponential divergence of the trajectories and the unpredictability of the time series. Subsequently, the surrogate data examination has been adopted to verify that the financial time series do not originate from a stochastic process. Lastly, a modified local nonlinear approximator has been presented to forecast the short-term behavior of the time series. The numerical simulations on the data collected from an interest-free economic institution in Iran have confirmed the complex nonlinear structure of the LPPs. The findings of this article may help the managers of the banks and economic organizations forecast the short-term horizon of the NPLs and, therefore, balance the budget and loans.

Data Availability

Access to data is restricted due to Omid Entrepreneurship Fund (OEF) policy. To access data, a request should be sent to OEF.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The Omid Entrepreneurship Fund of Iran financially supported this research.

References

[1] Y. Zhang, G. Yu, and D. Yang, "Predicting non-performing loan of business bank by multiple classifier fusion algorithms," *Journal of Interdisciplinary Mathematics*, vol. 19, no. 4, pp. 657–667, 2016.

[2] A. Belgrave, G. Kester, and M. Jackman, "Industry specific shocks and non-performing loans in Barbados," *the review of finance and banking*, vol. 4, no. 2, 2012.

[3] Y. Tang, H. Chen, B. Wang, M. Chen, M. Chen, and X. Yang, "Discriminant analysis of zero recovery for China's NPL," *Advances in Decision Sciences*, vol. 2009, Article ID 594793, 16 pages, 2009.

[4] S. L. Lin, "A new two-stage hybrid approach of credit risk in banking industry," *Expert Systems with Applications*, vol. 36, no. 4, pp. 8333–8341, 2009.

[5] A. O. Adewusi, T. B. Oyedokun, and M. O. Bello, "Application of artificial neural network to loan recovery prediction," *International Journal of Housing Markets and Analysis*, vol. 9, no. 2, pp. 222–238, 2016.

[6] G. Kou, Y. Peng, and C. Lu, "MCDM approach to evaluating bank loan default models," *Technological and Economic Development of Economy*, vol. 20, no. 2, pp. 292–311, 2014.

[7] P. Saha, I. Bose, and A. Mahanti, "A knowledge based scheme for risk assessment in loan processing by banks," *Decision Support Systems*, vol. 84, pp. 78–88, 2016.

[8] Z. Bitvai and T. Cohn, "Predicting peer-to-peer loan rates using bayesian non-linear regression," in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pp. 2203–2209, Austin, TX, USA, January 2015.

[9] R. Calabrese, "Predicting bank loan recovery rates with a mixed continuous-discrete model," *Applied Stochastic Models in Business and Industry*, vol. 30, no. 2, pp. 99–114, 2014.

[10] P. Tahmasebi and A. Hezarkhani, "A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation," *Computers & Geosciences*, vol. 42, pp. 18–27, 2012.

[11] B. LeBaron, "Chaos and nonlinear forecastability in economics and finance," *Philosophical Transactions of the Royal Society of London, Series A: Physical and Engineering Sciences*, vol. 348, no. 1688, pp. 397–404, 1994.

[12] M. Pourmahmood Aghababa and J. Abdollahi Sharif, "Chaos and complexity in mine grade distribution series detected by nonlinear approaches," *Complexity*, vol. 21, no. S2, pp. 355–369, 2016.

[13] S. Ramdani, F. Bouchara, and O. Caron, "Detecting high-dimensional determinism in time series with application to human movement data," *Nonlinear Analysis: Real World Applications*, vol. 13, no. 4, pp. 1891–1903, 2012.

[14] H.-L. Yang and H.-C. Lin, "Applying the hybrid model of EMD, PSR, and ELM to exchange rates forecasting," *Computational Economics*, vol. 49, no. 1, pp. 99–116, 2017.

[15] A. Papana, C. Kyrtsou, D. Kugiumtzis, and C. Diks, "Detecting causality in non-stationary time series using partial symbolic transfer entropy: evidence in financial data," *Computational Economics*, vol. 47, no. 3, pp. 341–365, 2016.

[16] V. K. Dabhi and S. Chaudhary, "Financial time series modeling and prediction using postfix-GP," *Computational Economics*, vol. 47, no. 2, pp. 219–253, 2016.

[17] P. Flaschel, F. Hartmann, C. Malikane, and C. R. Proaño, "A behavioral macroeconomic model of exchange rate fluctuations with complex market expectations formation," *Computational Economics*, vol. 45, no. 4, pp. 669–691, 2015.

[18] T. Chen, X. Li, and J. He, "Complex dynamics of credit risk contagion with time-delay and correlated noises," *Abstract and Applied Analysis*, vol. 2014, Article ID 456764, 10 pages, 2014.

[19] T. Chen, J. He, and J. Wang, "Bifurcation and chaotic behavior of credit risk contagion based on fitzhugh-nagumo system," *International Journal of Bifurcation and Chaos*, vol. 23, no. 07, Article ID 1350117, 2013.

[20] S. Lahmiri, "On fractality and chaos in Moroccan family business stock returns and volatility," *Physica A: Statistical Mechanics and Its Applications*, vol. 473, pp. 29–39, 2017.
[21] S. Radhakrishnan, A. Duvvuru, S. Sultornsanee, and S. Kamarthi, "Phase synchronization based minimum spanning trees for analysis of financial time series with nonlinear correlations," *Physica A: Statistical Mechanics and Its Applications*, vol. 444, pp. 259–270, 2016.
[22] Q. Tian, P. Shang, and G. Feng, "The similarity analysis of financial stocks based on information clustering," *Nonlinear Dynamics*, vol. 85, no. 4, pp. 2635–2652, 2016.
[23] B. B. Nair, P. K. S. Kumar, N. R. Sakthivel, and U. Vipin, "Clustering stock price time series data to generate stock trading recommendations: an empirical study," *Expert Systems with Applications*, vol. 70, pp. 20–36, 2017.
[24] A. Parida, R. Bisoi, P. Dash, and S. Mishra, "Financial time series prediction using a hybrid functional link fuzzy neural network trained by adaptive unscented kalman filter," in *Proceedings of the 2015 IEEE Power, Communication and Information Technology Conference (PCITC)*, pp. 568–575, Bhubaneswar, India, 15 October 2015.
[25] R. Singh and S. Srivastava, "Stock prediction using deep learning," *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 18569–18584, 2017.
[26] T. Zhou, S. Gao, J. Wang, C. Chu, Y. Todo, and Z. Tang, "Financial time series prediction using a dendritic neuron model," *Knowledge-Based Systems*, vol. 105, pp. 214–224, 2016.
[27] R. Chandra and S. Chand, "Evaluation of co-evolutionary neural network architectures for time series prediction with mobile application in finance," *Applied Soft Computing*, vol. 49, pp. 462–473, 2016.
[28] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*, pp. 366–381, Springer, Warwick, England, 1981.
[29] T. A. Alexeeva, N. V. Kuznetsov, and T. N. Mokaev, "Study of irregular dynamics in an economic model: attractor localization and Lyapunov exponents," *Chaos, Solitons & Fractals*, vol. 152, Article ID 111365, 2021.
[30] N. V. Kuznetsov, T. N. Mokaev, O. A. Kuznetsova, and E. V. Kudryashova, "The Lorenz system: hidden boundary of practical stability and the Lyapunov dimension," *Nonlinear Dynamics*, vol. 102, no. 2, pp. 713–732, 2020.
[31] N. V. Kuznetsov, G. A. Leonov, T. N. Mokaev, A. Prasad, and M. D. Shrimali, "Finite-time Lyapunov dimension and hidden attractor of the Rabinovich system," *Nonlinear Dynamics*, vol. 92, no. 2, pp. 267–285, 2018.
[32] J. Holzfuss and G. Mayer-Kress, "An approach to error-estimation in the application of dimension algorithms," in *Dimensions and Entropies in Chaotic Systems*, pp. 114–122, Springer, New York, NY, USA, 1986.
[33] W. Liebert and H. Schuster, "Proper choice of the time delay for the analysis of chaotic time series," *Physics Letters A*, vol. 142, no. 2-3, pp. 107–111, 1989.
[34] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, no. 2, pp. 1134–1140, 1986.
[35] K. P. Harikrishnan, R. Misra, G. Ambika, and A. K. Kembhavi, "A non-subjective approach to the GP algorithm for analysing noisy time series," *Physica D: Nonlinear Phenomena*, vol. 215, no. 2, pp. 137–145, 2006.
[36] D. S. Broomhead and G. P. King, "Extracting qualitative dynamics from experimental data," *Physica D: Nonlinear Phenomena*, vol. 20, no. 2-3, pp. 217–236, 1986.
[37] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical Review A*, vol. 45, no. 6, pp. 3403–3411, 1992.
[38] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D: Nonlinear Phenomena*, vol. 110, no. 1-2, pp. 43–50, 1997.
[39] S. Mehdizadeh, "A robust method to estimate the largest Lyapunov exponent of noisy signals: a revision to the Rosenstein's algorithm," *Journal of Biomechanics*, vol. 85, pp. 84–91, 2019.
[40] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, "Determining Lyapunov exponents from a time series," *Physica D: Nonlinear Phenomena*, vol. 16, no. 3, pp. 285–317, 1985.
[41] M. T. Rosenstein, J. J. Collins, and C. J. De Luca, "A practical method for calculating largest Lyapunov exponents from small data sets," *Physica D: Nonlinear Phenomena*, vol. 65, no. 1-2, pp. 117–134, 1993.
[42] T. Schreiber and A. Schmitz, "Improved surrogate data for nonlinearity tests," *Physical Review Letters*, vol. 77, no. 4, pp. 635–638, 1996.
[43] C. T. Dhanya and D. Nagesh Kumar, "Nonlinear ensemble prediction of chaotic daily rainfall," *Advances in Water Resources*, vol. 33, no. 3, pp. 327–347, 2010.
[44] H. Abarbanel, *Analysis of Observed Chaotic Data*, Springer Science & Business Media, Berlin, Germany, 2012.
[45] K.-I. Itoh, "A method for predicting chaotic time-series with outliers," *Electronics and Communications in Japan*, vol. 78, no. 5, pp. 44–53, 1995.
[46] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, England, 2004.