“The use of chaos theory predicting the EURIBOR index”


| | |
|---|---|
| **AUTHORS** | Mike P. Hanias<br>Panayiotis G. Curtis<br>Ioannis Mylonakis |


| | |
|---|---|
| **ARTICLE INFO** | Mike P. Hanias, Panayiotis G. Curtis and Ioannis Mylonakis (2008). The use of chaos theory predicting the EURIBOR index. *Investment Management and Financial Innovations*, *5*(4-1) |
| **RELEASED ON** | Thursday, 15 January 2009 |
| **JOURNAL** | "Investment Management and Financial Innovations" |
| **FOUNDER** | LLC “Consulting Publishing Company “Business Perspectives” |

Mike P. Hanias (Greece), Panayiotis G. Curtis (Greece), Ioannis Mylonakis (Greece)

# The use of chaos theory predicting the EURIBOR index


## Abstract

In this paper the chaos theory is used to predict the EURIBOR time series index from the reconstruction of its attractor. A non linear time series technique is applied using data of one week EURIBOR rate. For this purpose the optimal delay time and the minimum embedding dimension using the method of False Nearest Neighbors were found. From reconstruction of the corresponding strange attractor a 30 time steps out of sample prediction of the EURIBOR index is achieved. It indicates that the specific method could be used to facilitate the decision making process (especially for investment purposes), which requires as an important input the future rate or EURIBOR, since it could predict it for more than 6 months ahead.


**Keywords:** interest rate forecasting, financial markets, nonlinear time series.
**JEL Classification:** C53, E4.

## Introduction

Chaos theory has been applied in a wide variety of fields, e.g. physics, engineering, ecology and economics. The economist interest in chaotic system is focus on the ability of forecasting such a time series. Chaotic systems are deterministic systems governed by a low number of variables which display a quite complex behavior. These systems are unpredictable in the long term due to their ability to amplify even a very small initial perturbation of initial conditions.

In this paper a non linear time series technique is applied, using EURIBOR rate for one week index data from 30-12-1998 to 9-11-2007, in order to characterize and predict the time series. The paper is organized in two steps. In a first step state space parameters as the time delay and embedding dimension have been obtained for the above mentioned time series in order to carry out their analysis in the reconstructed state space. In a second step out of sample time series prediction is achieved using the reconstructed state space.

## 1. Time series

The EURIBOR time series index is presented as a signal $x = x(t)$ as shown in Figure 1. It covers data from 30-01-1998 to 9-11-2007, representing the official EURIBOR rate for one week time. The sampling rate was $\Delta t$ = 1 week and the number of data are N = 2277.

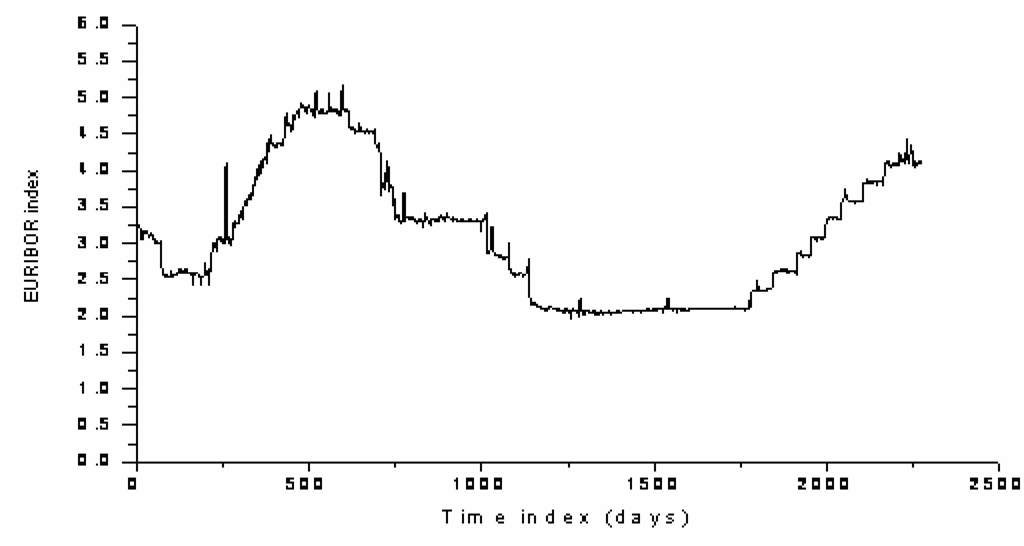


**Fig. 1. Time series of EURIBOR index**

## 2. State space reconstruction

From our data we construct a vector $y(t(i))$, $i$ = 1 to $N$, in the $m$ dimensional phase space given by the following relation (Kantz & Schreiber, 1997; Takens, 1981)

$$y(t(i)) = (x(t(i)), x(t(i+\tau)), \ldots, x(t(i+(m-2)\tau)), x(t(i+(m-1)\tau))). \quad (1)$$

This vector represents a point to the $m$ dimensional phase space in which the attractor is embedded each time, where $\tau$ is the time delay $\tau = i\Delta t$ while $\Delta t$ = 1 day. The term $x(t(i))$ represents a value of the examined scalar time series in time, corresponding to the $i$-th component of the time series. Use of this method reduces phase space reconstruction to the problem of proper determining suitable values of $m$

and $\tau$. The choice of these values is not always simple, especially when we do not have any additional information about the original system and the only source of data is a simple sequence of scalar values, acquired from the original system. The dimension, where a time delay reconstruction of the phase space provides a necessary number of coordinates (Strozzi, 2002) is called embedding dimension $m$.

**2.1. Time delay *τ*.** Using the average mutual information we can obtain $\tau$, less associated with linear point of view, and thus more suitable for dealing with nonlinear problems.

The average mutual information may be expressed by the following formula (Fraser & Swinney, 1986, Abarbanel; 1996):

$$I(\tau) = \sum_{x(t(i)),x(t(i+\tau))} P(x(t(i)), x(t(i+\tau))) log_2 \left( \frac{P(x(t(i)), x(t(i+\tau)))}{P(x(t(i)))P(x(t(i+\tau)))} \right), \quad (2)$$

where $P(x(t(i)), x(t(i+\tau)))$ is the joint probability density for the values $x(t(i))$ and $x(t(i+\tau))$, while $P(x(\tau))$ expresses the probability density of the value $x(\tau(i))$. In general, $I(\tau)$ expresses the amount of information (in bits), which may be extracted from the value in time $t(i)$ about the value in time $t(i+\tau)$. A value of τ, suitable for the phase space reconstruction, is usually considered the position of the first minimum of $I(\tau)$ (Kantz & Schreiber, 1997). In this case $\tau = 48$ time steps as shown in Figure 2.

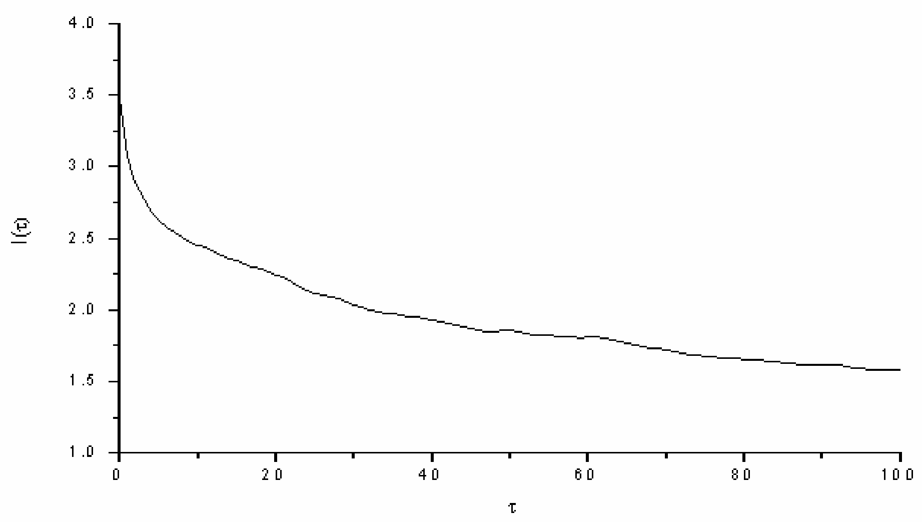


**Fig. 2. Mutual information *I* vs time delay τ**

**2.2. Embedding dimension *m*.** After obtaining the satisfactory value of $\tau$, the embedding dimension $m$ is to be determined in order to finish the phase space reconstruction. For this purpose the method of False Nearest Neighbors (Kennel et al., 1992) is used. More specifically, the method is based on a fact that when embedding dimension is too low, the trajectory in the phase space will cross itself. If we are able to detect these crossings, we may decide whether the used $m$ is large enough for correct reconstruction of the original phase space (i.e. when no intersections occur) or not. When intersections are present for a given $m$, the embedding dimension is too low and we have to increase it at least by one. Then, we test the eventual presence of self-crossings again (Kennel et al., 1992, Abarbanel, 1996). The practical realization of the described method is based on testing the neighboring points in $m$-dimensional phase space. Typically, certain amount of points is taken in the phase space and finds the nearest neighbor to each of them. Then distances for all these pairs are computed as well as their distances in $(m+1)$-dimensional phase space. The rate of these distances is given by

$$P = \frac{\| y_i(m+1) - y_{n(i)}(m+1) \|}{\| y_i(m) - y_{n(i)}(m) \|}, \quad (3)$$

where $y_i(m)$ represents the reconstructed vector, belonging to the $i$-th point in the $m$-dimensional phase space and index $n(i)$ denotes the nearest neighbor to the $i$-th point. If $P$ is greater than some value $P_{max}$, we call this pair of points false nearest neighbors (i.e. neighbors, which arise from trajectory self-intersection and not from the closeness in the original phase space). In the ideal case, when the number of false neighbors falls to zero, then the value of $m$ is found. For this purpose the rate of false nearest neighbors is computed in the reconstructed phase space using the formula

$$\left| x_{i+m\tau} - x_{n(i)+m\tau} \right| \geq R_A, \quad (4)$$

where $R_A$ is the radius of the attractor,

$$R_A = \frac{1}{N}\sum_{i=1}^{N}|x_i - \overline{x}| \quad (5)$$

and

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (6)$$

is the average value of time series.

When the following criterion

$$P \geq P_{max}, \quad (7)$$

is satisfied then it can be used to distinguish between true and false neighbors (Abarbanel, 1996). The dimension $m$ is found when the percent of false nearest neighbors decreases below some limit, typically set to 1% (Kugiumtzis et al., 1994), thus $P_{max}$=10 is chosen. Matlab code is used to calculate the mutual information $I$ and the quantity $P$. Figure 3 shows the situation for the system (the percent of false nearest neighbors number vs. total neighbors number is displayed). The percentage of false neighbors that is under the above limit is achieved for $m$ = 6.

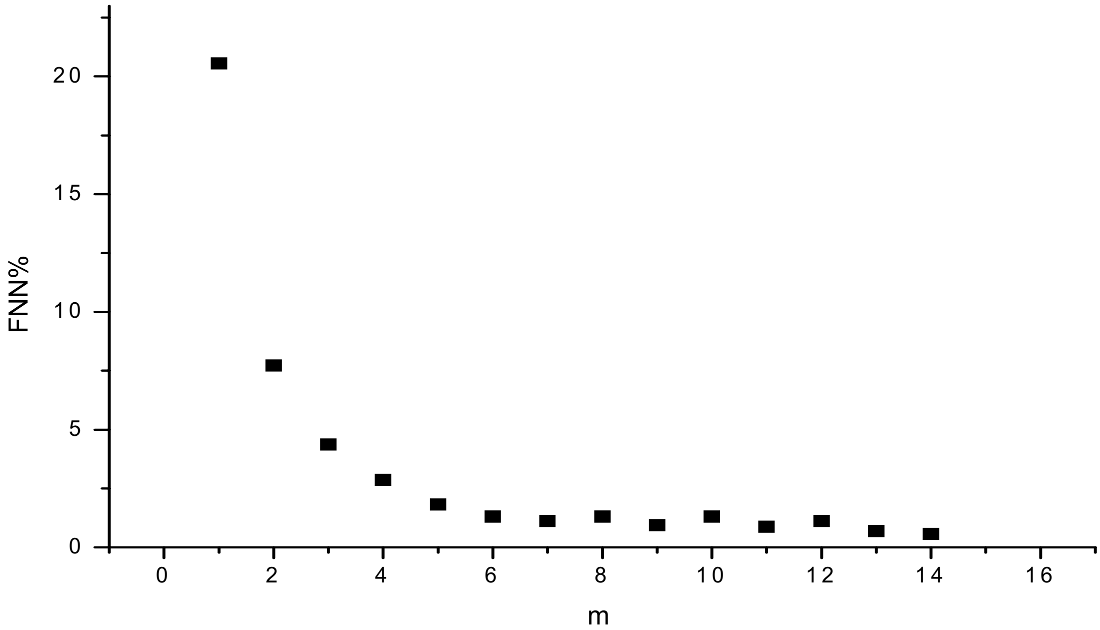


**Fig. 3. Percent of false nearest neighbor's number % FNN vs *m***

## 3. Time series prediction

The next step is to predict evolution of the examined quantity by computing weighted average of evolution of close neighbors of the predicted state in the reconstructed phase space (Miksovsky & Raidl, 2001; Stam et al., 1998; Hanias et al., 2007). The algorithm described by Sugihara and May for nonlinear forecasting with small modifications is used. Given a starting point in the time series $x(t)$ we would like to predict $x(t + 1)$, $x(t + 2)$, $x(t + 3)$ etc. a number of steps ahead, and compare the predictions, which we will designate $P_x(t + n)$, with the actual time series. Now for each vector $y(i)$ we located the $m$ + 1 nearest neighbor in the $m$-dimensional state space. We will designate the $k^{th}$ nearest neighbor vectors of $y(i)$ as $NN_{k,j}$. The $k$ index indicates the number (from 1 to $m$ + 1) of the nearest neighbor; the $j$ index is the time index in the original time series. We excluded nearest neighbors with time indices $j$ when $[i - j] < 3\tau$ autocorrelation time. This procedure is called "within-sample" prediction. Sugihara and May used "out-of-sample" prediction. For a time series of length $N$, out-of-sample prediction requires $i > 0.5\ x\ N$ and $j < 0.5\ x\ N+$ constant. There are no fundamental differences between the two procedures, only within sample prediction may be more suitable for short data sets. Now the predicted value for $n$ steps ahead prediction was given by

$$P_{x(i+n)} = \sum_{k=1}^{m+1} y(k_j + n) w_k \ . \quad (8)$$

Here

$$w_k = \frac{|y(i) - NN_{k,j}|^{-2}}{\sum_{k=1}^{m+1}|y(i) - NN_{k,j}|^{-2}} . \quad (9)$$

In our case we choose $m$ = 6, $\tau$ = 48 from our previous analysis and we put $k$ = 5 the number of neighbors which specifies number of points (nearest neighbors of the state the prediction is done from) that are used for the prediction and $n$ = 7, 15, 30 the number of steps forward – the prediction is done by this number of steps ahead. Figure 4 presents the prediction for $n$ = 7 days ahead.

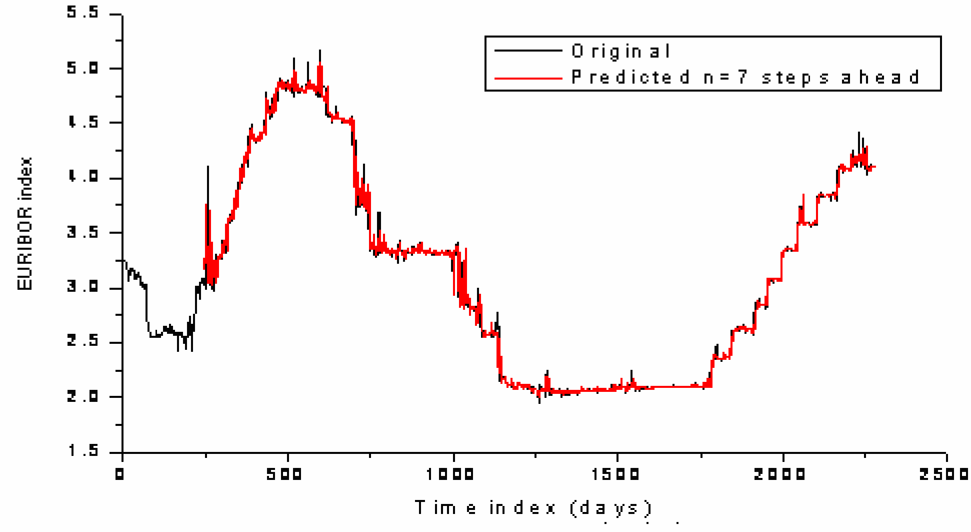


**Fig. 4. Actual and predicted time series for *n* = 7 time steps ahead**

Figure 5 presents the prediction for *n* = 15 days ahead.

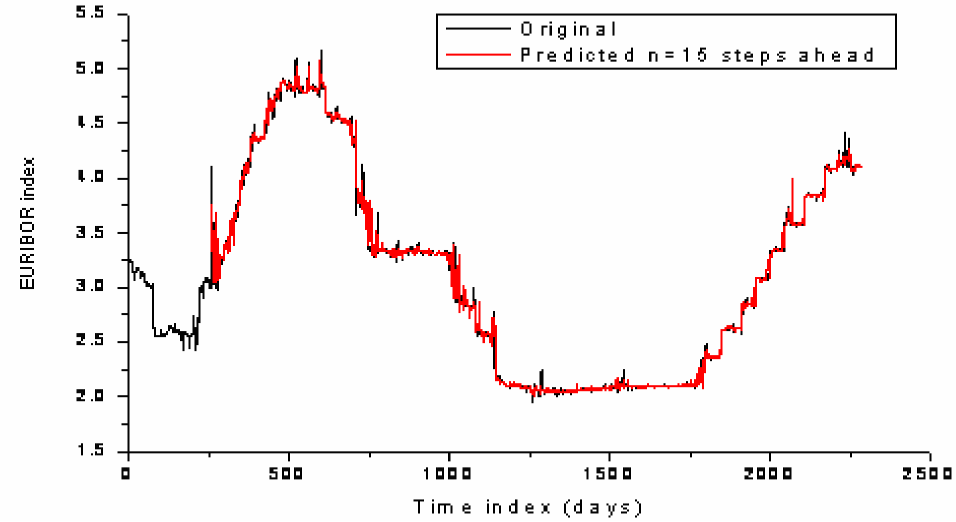


**Fig. 5. Actual and predicted time series for *n* = 15 time steps ahead**

Figure 6 presents the prediction for *n* = 30 days ahead.

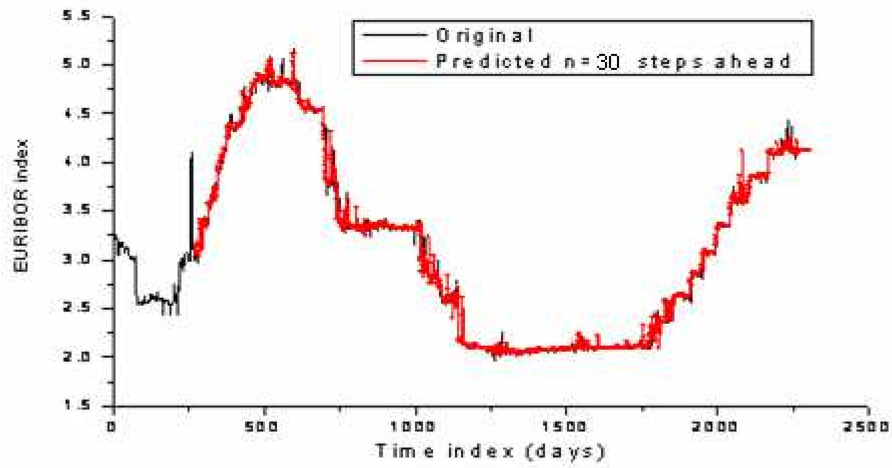


**Fig. 6. Actual and predicted time series for *n* = 30 time steps ahead**

Figures 4, 5 and 6 exhibit that after 30 steps ahead the predicted line is distancing itself substantially from the original one, showing that predictions under this precise method are not so valid.

Table 1 presents the actual predicted values for $n = 30$.

Table 1. Actual values and predicted out of sample values

| Actual values and predicted out of sample values for $n$ = 30 | |
|---|---|
| **Actual values** | **Predicted values** |
| 4.137 | 4.11048 |
| 4.101 | 4.13805 |
| 4.107 | 4.14463 |
| 4.089 | 4.13312 |
| 4.088 | 4.12699 |
| 4.115 | 4.08596 |
| 4.119 | 4.13466 |
| 4.136 | 4.1308 |
| 4.082 | 4.09873 |
| 4.036 | 4.09817 |
| 4.051 | 4.09535 |
| 4.081 | 4.07331 |
| 4.095 | 4.10501 |
| 4.097 | 4.09253 |
| 4.093 | 4.10618 |
| 4.09 | 4.18814 |
| 4.099 | 4.2219 |
| 4.113 | 4.08508 |
| 4.117 | 4.1072 |
| 4.113 | 4.11989 |
| 4.117 | 4.11817 |
| 4.125 | 4.10958 |
| 4.123 | 4.10748 |
| 4.12 | 4.10988 |
| 4.109 | 4.10622 |
| 4.103 | 4.11178 |
| 4.102 | 4.10973 |
| 4.101 | 4.10759 |
| 4.102 | 4.10726 |
| 4.103 | 4.11021 |
| 4.099 | 4.12087 |
| Until 2227 real data | |
| N/A | 4.11954 |
| N/A | 4.11302 |
| N/A | 4.11364 |
| N/A | 4.11368 |
| N/A | 4.11351 |
| N/A | 4.11352 |
| N/A | 4.11359 |

## Conclusion

In this chaotic analysis, non-linearity was discovered in EURIBOR data, and the analysis presented here examines this question further. The minimum embedding dimension is estimated to be 6. This means that the system is a high dimension chaotic system. From reconstruction of the system's strange attractor we achieved a 30 time steps out of sample prediction of the EURIBOR index.

## References


1. Abarbanel, H.D.I. (1996), "Analysis of observed chaotic data", Springer, New York, p. 272.
2. Fraser, A.M., Swinney H.L. (1986), "Independent coordinates for strange attractors from mutual information", Phys. Rev. A, 33, p. 1134.
3. Hanias, M.P., P.G, Curtis, J.E. Thallasinos (2007), "Non-Linear Dynamics and Chaos: The case of the Price Indicator at the Athens Stock Exchange", International Research Journal of Finance and Economics, Issue 11, pp. 154-163.
4. Kennel M.B., Brown R., Abarbanel H.D.I. (1992), "Determining embedding dimension for phase-space reconstruction using a geometrical construction", Phys. Rev. A, 45, p. 3403.
5. Kantz H. and T. Schreiber (1997), "Nonlinear Time Series Analysis", Cambridge University Press, Cambridge.
6. Kugiumtzis D., Lillekjendlie B., Christophersen N. (1994), "Chaotic time series", Part I, Modeling Identification and Control 15, p. 205.
7. Miksovsky J., Raidl A. (2001), "On some nonlinear methods of meteorological time series analysis", WDS 2001 Conference.
8. Stam C.J., Pijn J.P.N., Pritchard W.S. (1998), "Reliable detection of nonlinearity in experimental time series with strong periodic component", Physica *D* 112, p. 361.
9. Strozzi F. (2002), "Application of Non Linear Time Series Using Exchange Rate", Liuc papers n. 99, Serie Metodi quantitative 14, Suppl a.gennaio.
10. Takens F. (1981), "Detecting strange attractors in turbulence", Lecture notes in Math., Vol. 898, Springer, New York.