APPLICATION OF NON-LINEAR TIME SERIES ANALYSIS TECHNIQUES TO HIGH FREQUENCY CURRENCY EXCHANGE DATA

Fernanda Strozzi

Contents

Introduction

Seminal work of Prigogine (Nicolis and Prigogine, 1977) and Haken (Haken, 1983) has lead to the realization that large classes of systems may exhibit abrupt transitions, hysteresis, spatio-temporal structures or deterministic chaos. This has questioned the reductionist paradigm, i.e. the reduction of observed phenomena to elementary entities at lower levels of hierarchy and organization. Furthermore, it has been observed that nonlinear phenomena, that are not adequately described by linear approximations, are encountered in all areas of science.

Despite this rich variety of nonlinear dynamical systems, there is accumulating evidence that certain complex scenarios are frequently repeated between different fields of science. These findings indicate, that although complex systems may differ substantially in their detailed properties, deep analogies on their organization and functioning exist. As a consequence there has been an increasing interest in the study of "complexity" and in the search of a common background to all these systems (Waldrop, 1993). The search of this common background has mainly concentrated in economics on two types of paradigms: Self Organized Criticality (SOC) and chaotic systems. SOC systems are deterministic nonequilibrium systems composed by many interacting parts which have the ability to develop structures and patterns in the absence of control or manipulation by an external agent (Jensen, 1998). This emergent behaviour, which the interacting parts cannot show alone, is not just the sum of their individual properties, and, although, it is dynamically complex, the statistical properties are described by simple power laws. SOC systems have been looked for in such diverse areas as geophysics (earthquakes), astrophysics (quasars), condensed matter physics, biological evolution and economics. The paradigm model for SOC is a sand pile.

The second class of complex systems are chaotic systems. Chaotic systems are deterministic systems governed by a "low" number of variables, which display a quite complex behaviour. Furthermore, even though chaotic systems are described by differential equations, which do not contain any random function, they are unpredictable in the long-term due to their ability to amplify even a very small initial perturbation of initial conditions.

Chaos theory have been also applied in a wide variety of fields, e.g. physics, chemistry, engineering, ecology and economics. The roots of economists' interest in chaotic systems are to be found in the vast non mathematical literature on business cycles. In fact, throughout the last century economists have postulated the existence of different dynamical behaviours in the form of economic cycles, including the business cycle (Mitchel, 1927), the Kuznets (Kuznets, 1973), and the Kondratieff cycle or economic long wave (Kondratieff, 1935). Since variations in amplitude and period have been observed it is clear that they are not regular cycles, but is there a manifestation of chaotic behaviour? Is it possible to find the degrees of freedom that govern such behaviour?

Specifically, in the last decades there have been a considerable amount of discussion relating the theory of Brownian motion (Osborne,1959; Malkiel, 1990), fractional Brownian motion (Mandelbrot, 1998), non-linearity (Brock *et al*., 1991), chaos and fractals (Hsieh, 1991; Lorenz, 1993; Peters, 1996), scaling beahviour (Mantegna and Stanley, 1995; Mantegna and Stanley, 1996), and self organized criticality (Bak and Chen,

1991; Shlesinger *et al*., 1993) to the study of financial time series. The problem of characterizing financial time series is still an open question. Most of the test developed in the area of economic theory, provide evidence of nonlinear dynamics, which is a necessary but not sufficient condition for chaos. This nonlinearity may be deterministic or not deterministic. In fact, there is no convincing evidence of deterministic low-dimensionality in price series (Scheinkman and LeBaron, 1989; Papaioannou and Karytinos, 1995) and the claims of low-dimensional chaos have never been well-justified. For example, Andreadis (2000) analysing the S&P 500 index time series favours the stochastic hypothesis, whereas Friederich *et al*. (2000) using the high frequency price changes of the US dollar-German Mark exchange rates supports the analogy of turbulence and financial data (Mantenga and Stanley, 1996).

In this work we have applied non-linear time series techniques to high frequency currency exchange data from the HFDF96 data set provided by Olsen & Associated. The time series studied are the exchange rates between the US dollar and 18 other foreign currencies from the Euro zone - i.e. Belgium Franc (BEF), Finnish Markka (FIM), German Mark (DEM), Spanish Peseta (ESP), French Franc (FRF), Italian Lira (ITL), Duch Guilder (NLG), and finally ECU (XEU)- and outside –Australian Dollar (AUD), Canadian Dollar (CAD), Swiss Franc (CHF), Danish Krone (DKK), British Pound (GBP), Malaysian Ringgit (MYR), Japanese Yen (JPY), Swedish Krona (SEK), Singapore dollar (SGD) and South African Rand (ZAR)-. Our main interest is on trying to classify these series and analysing if their dynamical behaviours are in some way correlated. Furthermore, we are also interested in comparing the Euro with the exchange currencies that are part of it to analyse its behaviour and properties in comparison with them.

First, a preliminary study has been carried out with the aim of characterising those time series in terms of power spectral distribution, long term memory (R/S analysis), and stationarity (Space-time separation plots).

In a second step state space reconstruction parameters, i.e. 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. The non existence of stationarity calls for the application of Recurrence Quantification Analysis (RQA). This method is based of the definition of several parameters that allow the quantification of the Recurrence Plots (RP) introduced by Eckmann *et al*. (1987). The recurrence plot is based on the computation of the distance matrix between the reconstructed points in the phase space. This produces an array of distances in a square matrix. In order to extend the original concept and made it more quanitative Zbilut and Webber (1992) developed a methodology called recurrence quantification analysis (RQA) (Webber and Zbilut, 1994). As a result, they defined several variables to quantify RPs, which are: *%recur* (percentage of neighbouring points in recurrence plot); *%deter* (percentage of recurrent points forming diagonal line structures); *entropy* (Shannon entropy of line segments distributions); *trend* (measure of the paling recurrence points away from the central diagonal); $1/line_{max}$ (reciprocal of the longest diagonal line segment which may relate directly to largest positive Lyapunov exponent)(Trulla *et al*., 1996).

The RQA analysis of each currency exchange rate time series have shown a certain coherent structure. This structure allows a preliminary classification of the time series in several clusters. Moreover, the RQA analysis was repeatedly performed on 336-point epochs in order to analyse the dynamic information obtained. Neighbouring epochs were shifted by 48 points and the nonlinear variables *%recur*, *%deter* and $line_{max}$ obtained for the 18 time series analysed. As discuss in the text it is possible to correlate certain events that

occurred during 1996 with those variables, for example, the entries of the Finnish Markka and the Italian Lira in the European Monetary System. Furthermore, the RQA method allows studying the degree of correlation between several data sets by comparing the evolution of the *%recur* in the time series. The analysis shows, as foreseen, the high correlation between the exchange currencies in the Euro zone and for how long correlation between several exchange currencies persist. Based on these results a new non-linear indicator of correlation between financial time series has been defined. This indicator shows, for example the high correlation between Japanese Yen and Canadian dollar and the fact that British pound was, in 1996, more correlated to those series than to those that form the Euro zone. The RQA methodology has shown its potential in analysing high frequency currency exchange rates and how it can be used to find prediction windows and correlation windows between several currencies. In fact RQA methodology is able to detect changes in currency exchange rates states. This would allow improving the forecasting capabilities by considering at the same time simultaneously correlated time series, as it has been shown for the case of prediction in nonlinear dynamical systems (Judd and Mees, 1998).

2. Data provision and treatment

The data sets used have been purchased by LIUC Library from Olsen & Associates. The data set used is called HFDF96 and is a subset of the O&A data bank that has been collected through real-time data-feeds using proprietary O&A data collection software. The data set HFDF96 consists of 43 intraday time series composed of:

1. Half hourly bid and ask quotes for 25 major foreign exchange spot rates:

USD/DEM	USD/JPY	USD/CHF	GPB/USD	USD/FRF
USD/ITL	USD/NLG	AUD/USD	CAD/USD	USD/SEK
USD/DKK	USD/FIM	USD/ZAR	USD/ESP	USD/BEF
USD/XEU	USD/SGD	USD/MYR	DEM/JPY	DEM/SEK
DEM/ESP	DEM/ITL	DEM/FIM	DEM/FRF	GPB/DEM

2. Half hourly bid and ask quotes for 4 precious metals spot prices (XAU; XAG; XPT; XPD)
3. 12 half hourly series of transaction prices form 6 major Euromarket future contracts (USD, DEM, GPB, CHF, ITL, XEU).
4. 2 stock indices: half hourly values of the Dow Jones Industrial Average and the S&P 500

The data span a period of 1 year from January 1996 to 31 December 1996. Each record includes the interpolation times (GMT).

In this work, we have analyzed half hourly bid ask quotes for some of the major foreign exchange spot rates. Table 1 summaries the foreign exchange rates considered.

Table 1. Data considered: Foreign exchange rates against US Dollar.

Data set	**Name**	**Data set**	**Name**
AUD	Australian dollar	GBP	British Pound
BEF	Belgium Franc	ITL	Italian Lira
CAD	Canadian dollar	MYR	Malaysian Ringgit
CHF	Swiss Franc	JPY	Japanese Yen
DEM	German Mark	NLG	Duch Guilder
DKK	Danish Krone	SEK	Swedish Krona
ESP	Spanish Peseta	SGD	Singapore Dollar
FIM	Finnish Markka	XEU	ECU
FRF	French Franc	ZAR	South African Rand

2.1. Data treatment

We will consider the logarithmic middle price y_m as our primary time series, which can be calculated as follows:

$$y_m = \frac{\log(p_{bid}) + \log(p_{ask})}{2} \qquad (1)$$

where p_{bid} and p_{ask} are the bid and ask prices of the US dollar with respect to some currency, respectively. In order to compare the different data sets analysed, we have normalised data sets between 0 and 1 and obtained a normalised logarithmic middle price y as follows:

$$y = \frac{y_m - \min(y_m)}{\max(y_m) - \min(y_m)} \qquad (2)$$

The foreign exchange spot rates AUD\USD, GPB\USD and CAD\USD have been inverted and the same transformation as defined above have been carried out. Figures 1-18 represent the data sets after transformation. . Each data set contains 17568 points, which corresponds to 48 points each day times 366 days, and hence, for example, 10 February is in correspondence with the period 1968-2016.

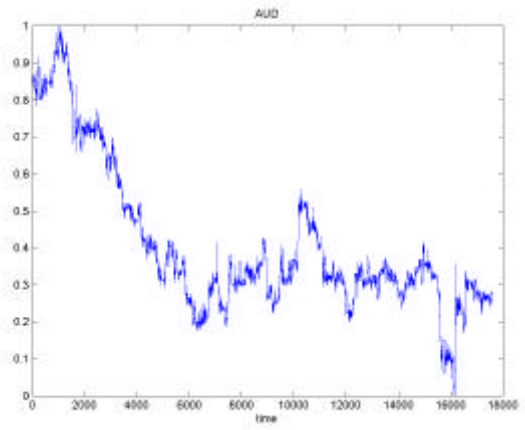

Figure 1. Australian-US dollar foreign exchange time series.

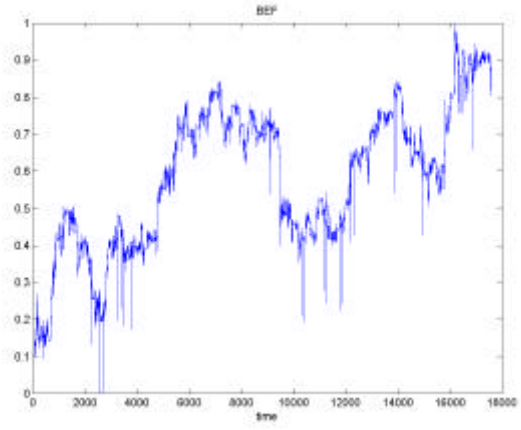

Figure 2. Belgium Franc-US dollar foreign exchange time series.

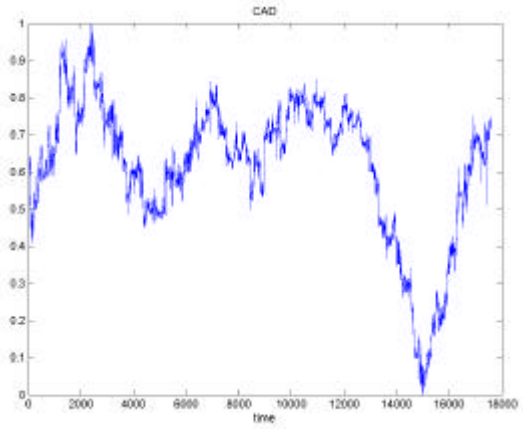

Figure 3. Canadian-US dollar foreign exchange time series.

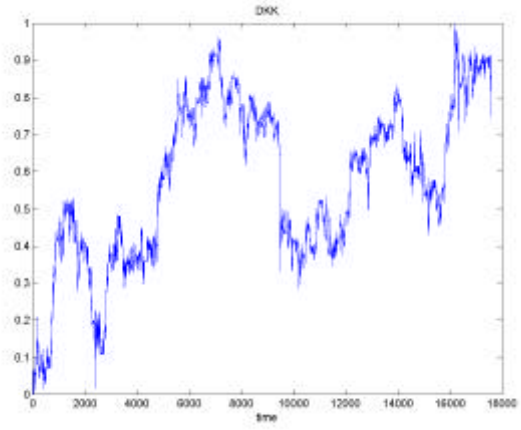

Figure 6. Danish Krone-US dollar foreign exchange time series.

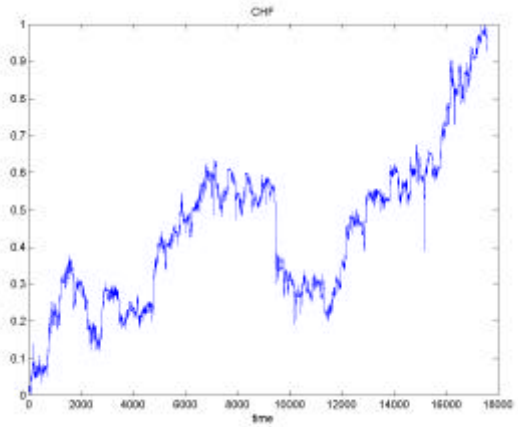

Figure 4. Swiss Franc-US dollar foreign exchange time series.

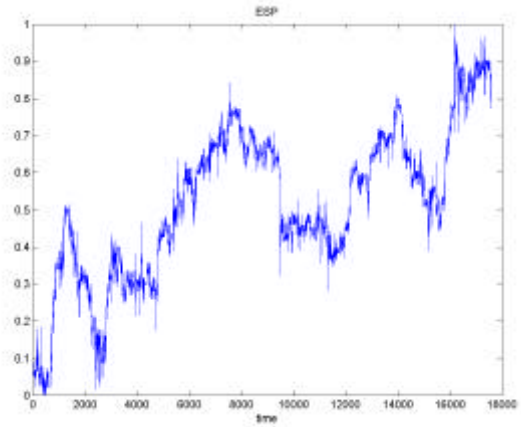

Figure 7. Spanish Peseta-US dollar foreign exchange time series.

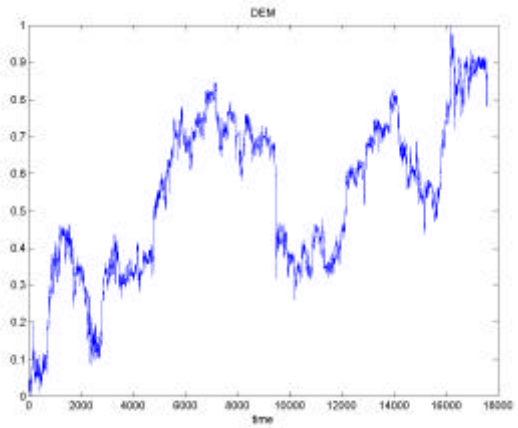

Figure 5. German Mark-US dollar foreign exchange time series.

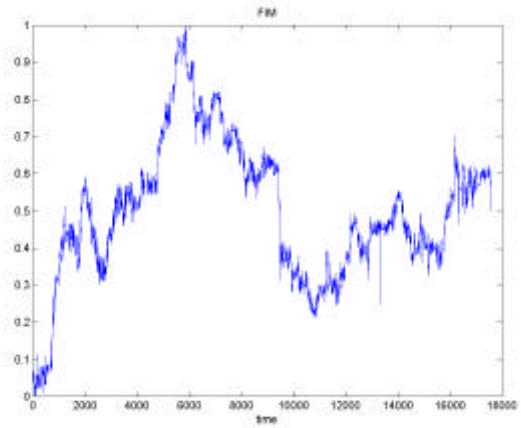

Figure 8. Finnish Markka-US dollar foreign exchange time series.

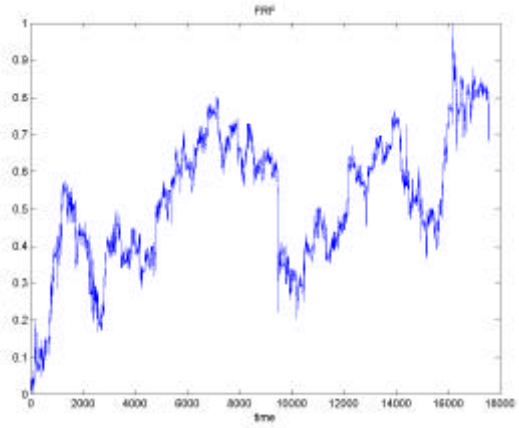

Figure 9. French Franc-US dollar foreign exchange time series.

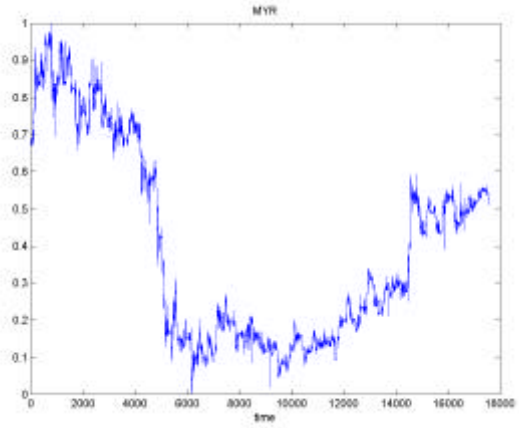

Figure 12. Malaysian Ringgit-US dollar foreign exchange time series.

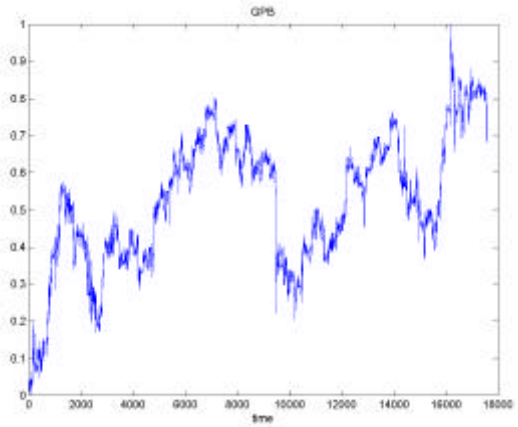

Figure 10. British Pound-US dollar foreign exchange time series.

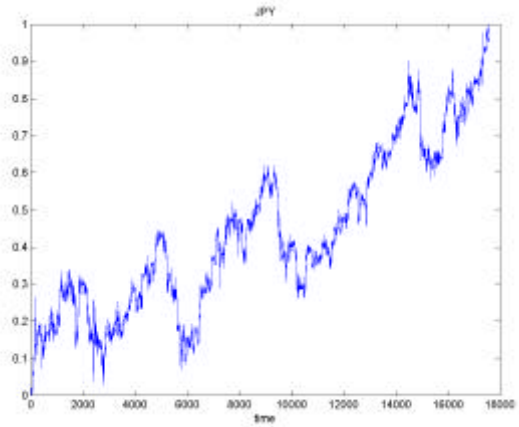

Figure 13. Japanese Yen-US dollar foreign exchange time series.

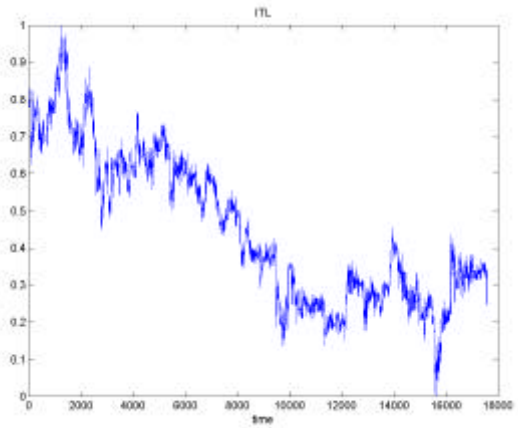

Figure 11. Italian Lira-US dollar foreign exchange time series.

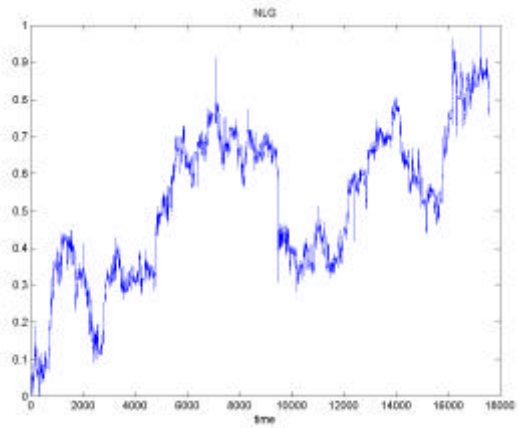

Figure 14. Duch Guilder-US dollar foreign exchange time series.

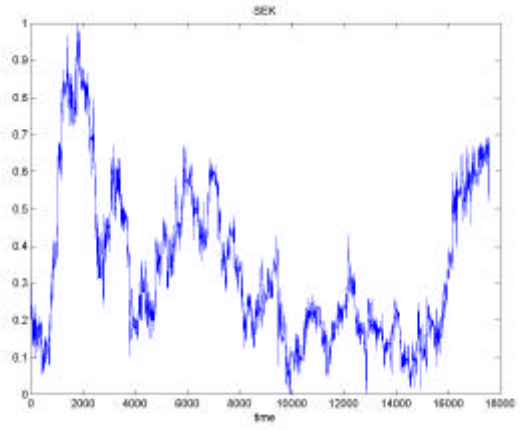

Figure 15. Swedish Krona-US dollar foreign exchange time series.

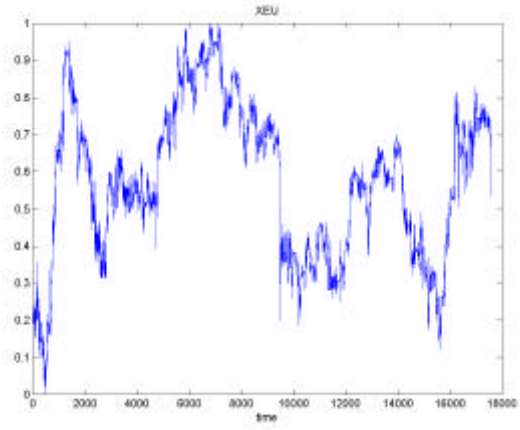

Figure 17. Euro-US dollar foreign exchange time series.

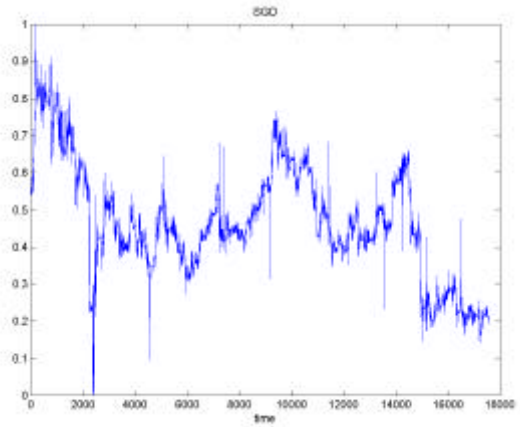

Figure 16. Singapore dollar-US dollar foreign exchange time series.

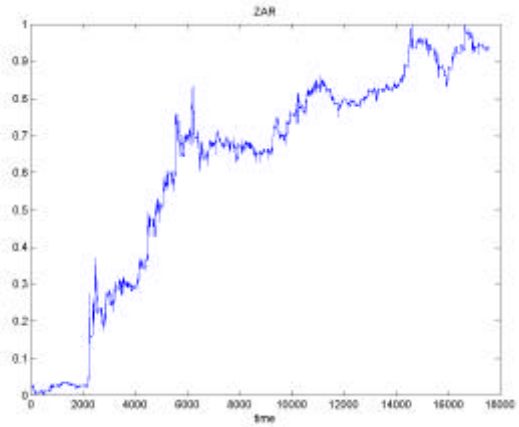

Figure 18. South African Rand-US dollar foreign exchange time series.

2.2. Historical background

Although the first stage in the Economic and Monetary Union process in the EU began on 1 July 1990 with the liberalisation of capital movements, the entry into force of the Treaty on European Union on 1 November 1993 marked the genuine starting-point of preparations for Economic and Monetary Union (EMU). In accordance with the Treaty, the second stage began on 1 January 1994 with, in particular, establishment of the European Monetary Institute (EMI), based at Frankfurt (Germany) as a first step towards the creation of the European System of Central Banks. The monetary turmoil experienced in 1995, largely caused by the slide in the value of the dollar, strengthened the Member States' political determination to go ahead with EMU. That determination took shape at the Madrid European Council of 15 and 16 December 1995, which confirmed that the third stage of Economic and Monetary Union was to go ahead on 1 January 1999 in accordance with the convergence criteria, the timetable, the protocols and procedures laid down in the Treaty. The single currency was decided to be called Euro.

Throughout 1996 and 1997 the economic upturn, against a background of closer nominal convergence, interest and inflation rates at exceptionally low levels, and stable exchange rates, enabled to be a general

improvement in the state of public finances, paving the way for the majority of Member States to switch to the Euro in 1999.

In 1991 Finland, Norway and Sweden related their currencies to the Euro (ECU). In 1992 due to the instabilities in the financial markets a 15 % fluctuation margin was conceded. Despite of that, the British pound as well as the Italian Lira abandoned the EMS in 1992. On 14 October 1996, the Finnish Markka joined the EMS Exchange-Rate Mechanism (ERM). The Italian Lira returned, on 25 November, to the ERM. Also in 1996 the Austrian Schilling joined the EMS. Only Greece, Sweden, Denmark and the United Kingdom were not members. From 1 January 1995 Austria, Finland and Sweden enter in the EU. From 1995 to 1997 the EU countries has to adequate their economical parameters to those defined in the Maastricht Treaty. On the 3 May 1998 the European Council confirmed the eleven countries that could adopt the Euro. However United Kingdom, Sweden and Denmark decided not to enter in the EMS. In this work we will not consider Portugal because the exchange rate were not available in HFDF96.

At the start of 1996 there were preoccupations that even those EU member states that were most enthusiastic about monetary union would have great difficulties in meeting the conditions for taking part in the planned move to a single currency in January 1999. The tough qualification criteria including limits on government budget deficits and government debt levels and the tide timetable made the financial markets sceptical. The attitude began to change after the meeting of the finance ministers of the EU governments in Verona (Italy), in April. There, it became clear that all member nations were determined to make the goal of the European Monetary Union (EMU) their economical and political priority. This was further confirmed when one after the other EU Member States announced strong austerity measures in the summer 1996 designed to reduce their budget deficits and meet the EMU criteria. Reflecting this remarkable political determination to achieve the single currency, the European financial markets gradually become less sceptical about the prospects for accomplishing it.

3. Embedding theory

The mathematical basis of continuous dynamical modelling is formed by differential equations of the following type:

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}, a) \qquad (3)$$

where the real variable t denotes time, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ represents the state variables of the system, depending on time t and on the initial conditions, and a_j are parameters of the system, while $\boldsymbol{F} = (F_1, F_2, \ldots, F_n)$ is a nonlinear function of these variables and parameters. Actual states of these systems are described by the vector variable $\boldsymbol{x}$ consisting of n independent components. Each state of the system corresponds to a definite point in phase space, which is called phase point. The time variation of the state of the system is represented as a motion along some curve called phase trajectory.

Experimentally, it is not always possible to measure the complete state of a system and, normally, when analysing a dynamical system, we have access to few observable quantities which, in the absence of noise, are related to the state space coordinates by:

$$s(t) = \mathbf{h}(\mathbf{x}(t)) \tag{4}$$

where ***h*** is normally an unknown nonlinear function called measurement function.

The theory of embedding is a way to move from a temporal time series of measurements to a state space "similar" -in a topological sense- to that of the underlying dynamical system we are interested in analysing. Techniques of state space reconstruction were introduced by Packard *et al.* (1981) and Takens (1981), who showed that it is possible to address this problem using measurements of a sufficiently long time series, *s(t)*, of the dynamical system of interest. Takens proved that, under certain conditions, the dynamics on the attractor of the underlying original system has a one-to-one correspondence with measurements of a limited number of variables. If the equations defining the underlying dynamical system are not known, and we are not able to measure all the state space variables, the state space of the original system is not directly accessible to us. However, if by measuring few variables we are able to reconstruct a one-to-one correspondence between the reconstructed state space and the original, this means that it is possible to identify unambiguously the original state space from measurements. This observation opened a new field of research called nonlinear time series analysis (Abarbanel, 1996; Diks, 1999, Kantz and Schreiber, 1997; Tong, 1990, amongst others). In the last few years, nonlinear time series analysis has expanded rapidly in the fields of Economics and Finance. Even though there is no conclusive evidence of chaotic structure, economic and financial time series seem to provide a promising area for the application of nonlinear approaches.

In order to explain the relationship that occurs between the reconstructed and the real state space, let us consider the following dynamical system

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}); \quad \mathbf{x} = (x_1, x_2, x_3) \tag{5}$$

We can define $\mathbf{y} = (y_1, y_2, y_3)$ as follows: $\mathbf{y} = (x_1, dx_1/dt, d^2x_1/dt^2)$, then the equations of motion take the form

$$\begin{aligned} \frac{dy_1}{dt} &= y_2 \\ \frac{dy_2}{dt} &= y_3 \\ \frac{dy_3}{dt} &= \mathbf{G}(y_1, y_2, y_3) \end{aligned} \tag{6}$$

for some function **G**. In this coordinate system, modelling the dynamics reduces to constructing the single function **G** of three variables, rather than three separate functions, each of three variables.

In this way we may proceed from the state space (x_1, x_2, x_3) to the space of derivatives $(x_1, dx_1/dt, d^2x_1/dt^2)$. The dynamics in this new space will be related to the dynamics of the original space by a nonlinear transformation, which is called the reconstruction map. The extension of this approach to higher-dimensional dynamical systems is straightforward by considering higher derivatives.

The advantage in considering the space of derivatives is that we can approximate them from measurements of x_1. But what kind of information about the original space is preserved in the new one?

There are two types of preserved information: qualitative and quantitative. Qualitative information is what allows a qualitative description of the dynamics described by topological invariants, such as for instance, singularity of the field, closeness of an orbit, stability of a fixed point, etc. (Gilmore, 1998) Quantitative information can be of two different types: geometrical and dynamical. Geometrical properties (Grassberger, 1983) consit of fractal dimensions or scaling functions. Dynamical methods (Wolf *et al*., 1985) rely on the estimation of local and global Lyapunov exponents and Lyapunov dimensions. In order to guarantee that the quantities computed for the reconstructed attractor are identical to those in the original state space, we require that the structure of the tangent space, i.e. the linearization of the dynamics at any point in the state space, is preserved by the reconstruction process. The problem is to see under what conditions this can happen.

Embedding theorems try to shed some light on this problem.

Let *s(t)* be the measure of some variable of our system, see Eq. (2). Takens (1981) shown that instead of derivatives, $\{s(t), \dot{s}(t), \ddot{s}(t), ...\}$, one can use delay coordinates, $\{s(t), s(t+\Delta t), s(t+2\Delta t), ...\}$, where *Dt* is a suitably chosen time delay. In fact, looking at the following approximation of the derivative of *s(t)*:

$$\frac{ds(t)}{dt} \cong \frac{s(t+\Delta t) - s(t)}{\Delta t} \tag{7}$$

$$\frac{d^2 s(t)}{dt^2} \cong \frac{s(t+2\Delta t) - 2s(t+\Delta t) + s(t)}{2\Delta t^2} \tag{8}$$

it is clear that the new information brought from every new derivative is contained in the series of the delay coordinates. The advantage of using delay coordinates instead of derivatives is that in case of high dimensions high order derivatives will tend to amplify considerably the noise in the measurements.

Another frequently used method, for state space reconstruction, is singular value decomposition (SVD), otherwise known as Karhunen-Loève decomposition, which was proposed by Broomhead and King (1986) in this context. The simplest way to implement this procedure is to compute the covariance matrix of the signal with itself and then to compute the eigenvalues, i.e. if *s(t)* is the signal at time *t*, the elements of the covariance matrix **Cov** are:

$$c_{ij} = \left\langle s(t)s(t+(i-j)t) \right\rangle^T \tag{9}$$

where *i* and *j* go from 1 to *n* where *n* is greater or equal to the dimension of the system in this new space. The eigenvectors of **Cov** define a new coordinate system. Typically, one calculates the dimension of the reconstructed phase space by considering only eigenvectors whose eigenvalues are "large". This method allows one to consider a time delay of one step and calculate an embedding dimension which is the rank of the covariance matrix **Cov**.

Then, from the space of derivatives, time lags or eigenvectors, it is possible to extract information about the underlying system, which was generating the measured data.

In order to preserve the structure of the tangent space and then the dynamic characteristic of it, the relation between the reconstructed space and the original one has to be an embedding of a compact smooth manifold

into R^{2n+1}, which means a one-to-one immersion i.e. a one-to-one C^1 map with Jacobian which has full rank everywhere. The point now is to show under what conditions the reconstruction forms an embedding.

A general existence theorem for embedding in Euclidean spaces was given by Whitney (1936) who proved that a smooth (C^2) n-dimensional manifold may be embedded in R^{2n+1}. This theorem is the basis of the time delay reconstruction (or embedding) techniques for phase space portraits from time series measurements proposed by Takens (1981), who proved that, under certain circumstances, if d_E -the dimension of the reconstructed state vector, normally called the embedding dimension- is greater or equal to *2n+1*, where *n* is the dimension of the original state space, then the reconstructed states fill out a reconstructed state space which is diffeomorphic, i.e. a one-to-one differentiable mapping with a differentiable inverse, to the original system. Generally speaking, the embedding dimension is the minimal number of dynamical variables with which we can describe the attractor when we know only one of its state variables or a function related to them.

Apart from the methods mentioned above, there are several other methods of reconstructing state space from the observed quantity s(t) that have appeared in the literature -for a critical review see Breeden and Packard (1994). Although the method of reconstruction can make a big difference in the quality of the resulting coordinates, it is not clear in general which method is the best. The lack of a unique solution for all cases is due in part to the presence of noise and to the finite length of the available data sets.

For Takens' theorem to be valid we need to assume that the underlying dynamics is deterministic and that both the dynamics and the observations are autonomous, i.e. **F** and **h** in Eqs. (3) and (4) depend only on $\boldsymbol{x}$ and not on *t*. Unfortunately, this is not the case of many systems in the field of control and communications which are designed to process some arbitrary input and hence, cannot be treated as autonomous. The extension of Takens' theorem to deterministically forced stochastic systems has been recently developed by Stark *et al*. (1987). In particular they proved that such an extension is possible for deterministically forced systems even when the forcing function is unknown, for input-output systems (which are just deterministic systems forced by an arbitrary input sequence) and for irregular sampled systems.

Another problem in embedding theory is that Takens' theorem has been proven for noise-free systems. Unfortunately, there is always a certain amount of noise, $s(t)$, in real data. Such noise can appear in both the measurements and the dynamics (Diks, 1999). Observational noise, i.e. $s(t)=h(\boldsymbol{x}(t))+s(t)$, does not affect the evolution of the dynamical system, whereas dynamical noise acts directly on the state of the dynamical system influencing its evolution, for example: $d\boldsymbol{x}/dt=\boldsymbol{F}(\boldsymbol{x},a)+s(t)$.

The effects of relatively small amount of observational noise may put severe restrictions on the characterisation and estimation of the properties of the underlying dynamical system. In order to remove the observational noise different possibilities are available and can be broadly divided into two categories: linear filters (Badii *et al*., 1998) and special nonlinear noise reduction methods that make use of the deterministic origin of the signal we are interested in (for a recent survey see: Kostelich and Schreiber, 1993; Davies, 1994). However, in the case of dynamical noise, the reconstruction theorem does not apply and it may even be impossible to reconstruct the state of the system (Takens,1996). In this situation, systems must be examined case by case before analysis. In particular, Stark *et al*. (1997) showed that the extension of Takens' theorem is possible for deterministic systems driven by some stochastic process.

3.1. Embedding parameters

The embedding theorem is important because it gives a rigorous justification for the state space reconstruction. However, Takens' theorem is true for the unrealistic case of an infinite, noise-free, number of points. Takens showed that, in this case, the choice of the time delay is not relevant, and gave indications only on the choice of the embedding dimension.

Nevertheless, in real applications, the proper choice of the time delay t and the calculation of an embedding dimension, d_E, are both fundamental for starting to analyse the data. As a matter of fact, a lot of research on state space reconstruction has centred on the problems of choosing the time delay and the embedding dimension which we can call the parameters of the reconstruction for delay coordinates.

If the time delay chosen is too small, there is almost no difference between the elements of the delay vectors, since all points are accumulated around the bisectrix of the embedding space: this is called redundancy (Casdagli *et al*., 1991). However, when t is very large, the different coordinates may be almost uncorrelated. In this case the reconstructed trajectory may become very complicated, even if the underlying "true" trajectory is simple: this is called irrelevance. Unfortunately no rigorous way exists of determining the optimal value of t. Moreover, similar problems are encountered for the embedding dimension. Working in a dimension larger than the minimum required by the data will lead to excessive requirements in terms of the number of data points and computation times necessary when investigating different questions such as, for example invariants calculation, prediction, etc. Furthermore, noise by definition has an infinite embedding dimension, so it will tend to occupy the additional dimensions of the embedding space where no real dynamics is operating and, hence, it will increase the error in the subsequent calculations. On the other hand, by selecting an embedding dimension lower than required, we would not be able to unfold the underlying dynamics, i.e. the calculations would be wrong since we do not have an embedding.

When derivatives, $\{s(t), \dot{s}(t), \ddot{s}(t), ...\}$, or SVD are employed there is no need to determine an optimum time delay. Nevertheless, for the case of derivatives, the reconstruction will depend on the way they are numerically calculated (which turns out to depend on different parameters, see for example (Burden and Faires, 1996) for a review of numerical calculation of derivatives). In practice for each method we will carry out a slightly different state space reconstruction. For the case of SVD, the time delay chosen is unitary, but there is still the problem of choosing the time scale or window in which the calculations are performed. Broomhead and King (1986) in fact, concluded that the effects of window length should be carefully investigated each time a state space reconstruction is carried out. Some details about the selection of time delay and embedding dimension are going to be discussed in great detail in the next Chapter.

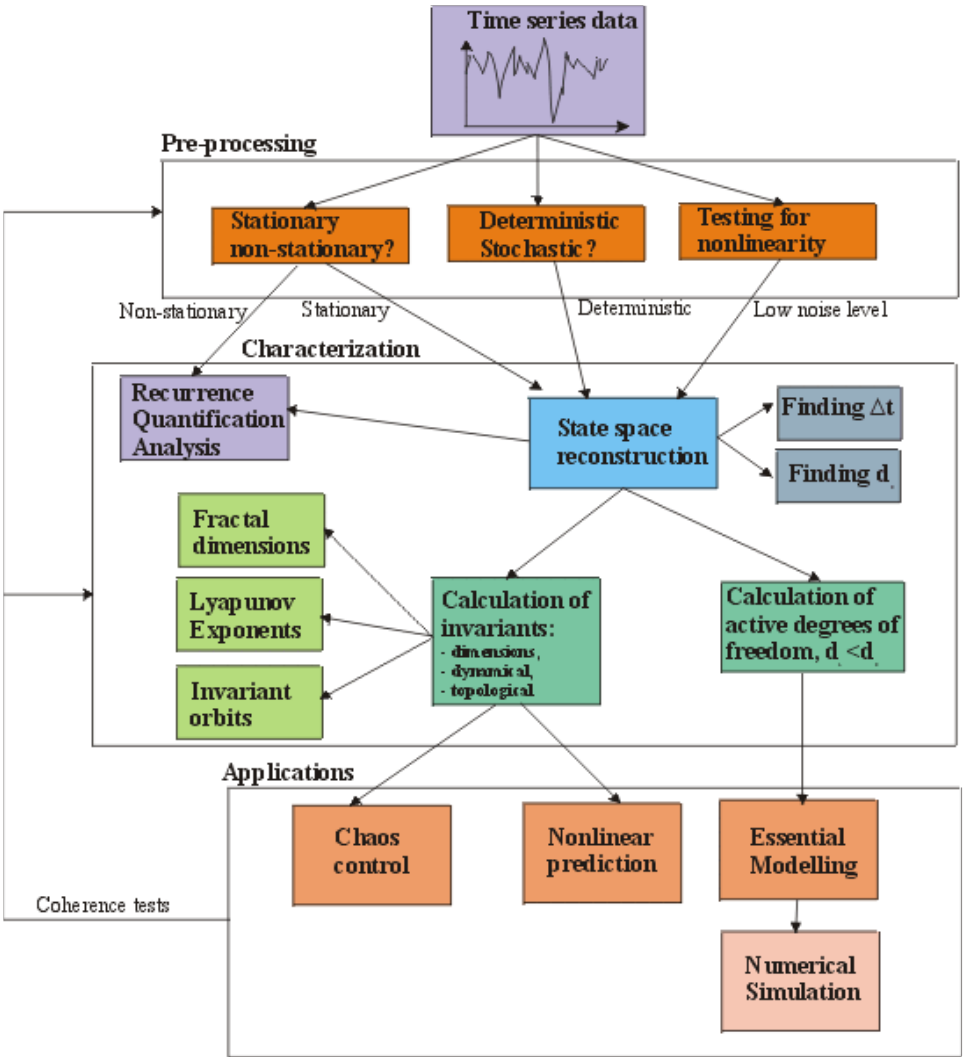

Figure 19. Schematic representation of nonlinear time series analysis using delay coordinate embedding (Strozzi and Zaldívar, 2001).

4. Chaotic time series analysis

Nonlinear analysis of experimental time series has, among its goals, the separation of high-dimensional and stochastic dynamics from low-dimensional deterministic signals, estimation of system parameters (characterisation), and , finally, prediction, modelling and control.

Unfortunately it seems very difficult to tell whether a series is stochastic or deterministically chaotic or some combination of these states. More generally, the extent to which a non-linear deterministic process retains its properties when corrupted by noise is also unclear. The noise can affect a system in different way, either in an additive way or as a measurement error, even though the equations of the system remain deterministic.

A schematic representation of the different steps is given in Fig. 19. Since a single reliable statistical test for chaoticity is not available, combining multiple tests is a crucial aspect, specially when one is dealing with limited and noisy data sets like in economic and financial time series.

There are different aspects that should be carefully studied before attempting to go further using nonlinear time series analysis methods. A long and exhaustive discussion can be found in Schreiber (1998) and the basic methodologies will be presented during the analysis part. Here, we shall briefly indicate the main problems one should be aware of. These can be summarized as follows:

- has the phenomenon been sufficiently sampled?;
- is the data set stationary or can one remove the nonstationary part?;
- is the level of noise sufficiently low so that one can obtain useful information using nonlinear time series techniques?

Some tests to study these questions have been recently implemented in the TISEAN software package (Hegger *et al*., 1999), which has incorporated a substantial quantity of algorithms developed for nonlinear time series analysis.

The problem of the number of samples needed to carry out state space reconstruction is related to the dimensionality of the problem we are dealing with. In order to characterize properly the underlying dynamics from the observed time series, we need to sample properly the phase space in which our dynamical system lies. As the dimension of the underlying system increases, a higher number of samples is needed. Ruelle (1990) discussed this problem, and based on simple geometrical considerations, he arrived at the following conclusion: if the calculated dimension of our system is well below $2log_{10}m$, where m is the total number of points in the original time series, then we are using a sufficient number of data points. Of course having a sufficient number of data points is a necessary but not a sufficient condition for reliable nonlinear time series analysis.

Another related problem is the sampling rate. Consider the case when we are sampling data from a, presumably, chaotic system. Chaotic systems, like stochastic ones, are unpredictable in the long run. This long run is related to the speed at which nearby trajectories diverge in phase space, which turns out to be related to the Lyapunov exponents of the system under study. Hence, if we are sampling at a rate slower than our predictability window, even though the underlying system is chaotic, we will find that our system behaves as a stochastic one. In this situation, if one suspects that the underlying system is deterministic, the best thing to do is to repeat the experiment by increasing the sampling rate. Interpolating between data points would be of no use as no new information is introduced.

A time series is said to be strictly stationary if its statistical distribution does not change across time. More specifically, suppose we have a set of m samples of the series $s(t)$ made at times t_1 through t_m, which need not be contiguous times. Strict stationarity implies that the joint probability density function of those m samples is identical to the joint probability distribution of other m samples taken at times t_{1+k} through t_{m+k}. This must be true for all the choices of m and k, as well for the m relative sample times. Why is stationarity so important? Because almost all methods developed by linear and nonlinear time series analysis assume that the time series we are analysing is stationary, which implies that the parameters of the system which has generated the time series, remain constant. For this reason time series analysis often requires one to turn a nonstationary series into a stationary one so as to use these theories. Unfortunately, nonstationary signals are very common in particular when observing natural or economical phenomena, and in some cases the nonstationary components, such as

the trend, may sometimes be of more interest than that of the stationary part obtained by removing the trend or the seasonal variation from the signal.

Even though a precise definition of stationarity exists, there is no magic formula for deciding whether a series is stationary or not. However, strong violations of the basic requirements that the dynamical properties of the system must not change, beyond their statistical fluctuations, can be checked simply by measuring such properties, i.e. mean, variance, spectral components, correlations, etc., for several segments of the data set. Nonlinear time series analysis has also developed its own techniques to study nonstationarity as we will see in Section 4.3.

4.1. Preliminary Analysis

4.1.1. R/S Analysis

A tool for studying long-term memory and fractality of a time series is the Rescaled Range analysis (R/S analysis) first introduced by Hurst (1951) in hydrology. Mandelbrot (1983) argued that R/S analysis is a more powerful tool in detecting long range dependence compared to more conventional analysis like autocorrelation analysis, variance ratios and spectral analysis. In this method, one measures how the range of cumulative deviations from the mean of the series is changing with the time. It has been found that, for some time series, the dependence of R/S on the number of data points (or time) follows an empirical power law described as $(R/S)_n=(R/S)_0\, n^H$, where $(R/S)_0$ is a constant, n is the time index for periods of different length, and H is the Hurst exponent. $(R/S)_n$ is defined as

$$\left(\frac{R}{S}\right)_n = \frac{\max_{1\le t\le n} A(t,n) - \min_{1\le t\le n} A(t,n)}{\sqrt{\frac{1}{n}\sum_{t=1}^{t}(s(t)-\langle s\rangle_n)^2}} \qquad (10)$$

where $A(t,n)$ is the accumulated departure of the time series $s(t)$ from the time average over the time interval n: $\langle s\rangle_n \quad A(t,n)=\sum_{i=t}^{t+n}(s(i)-\langle s\rangle_n)$.

The Hurst exponent, $0 \pounds H \pounds 1$, is equal to 0.5 for random, white noise series, <0.5 for rough anticorrelated series, and >0.5 for positively correlated series. Financial time series have been found to exhibit some universal characteristics that resemble the scaling laws typical of natural systems in which large numbers of units interact. For instance, the Hurst exponent has been extensively applied by Peters (1996) to various capital markets and in most of the cases he has found persistent memory, i.e. fractal structure and nonperiodic cycles.

In this work we have used the scaled windowed variance method (Cannon *et al*., 1996) to estimate H by linear regression of $log(R/S)$ versus $log(Windowsize)$. According to this method the signal is repeatedly divided into windows, but instead of computing the standard deviation of the means within the windows, the means of the standard deviations within the windows are used to obtain an estimate of H. The results for two time series are presented in Figs. 20-21, whereas the H values are summarised in Table 2. As it can be seen the currency exchange rates time series show long memory effects, $H \gg 0.9$ for all time series studied, which indicates that these series are persistent and fractal with long-term memory. This value might be consistent with either some

kind of fractal noise. Nevertheless with the Hurst exponent it is not possible to differentiate an exchange rate time series from another.

Table 2. Estimated *H*, using the standard scaled window variance method for the foreign exchange time series.

Data set	H	Data set	H
AUD	0.991	GBP	0.992
BEF	0.989	ITL	0.995
CAD	0.994	MYR	0.991
CHF	0.991	JPY	0.991
DEM	0.989	NLG	0.990
DKK	0.991	SEK	0.988
ESP	0.991	SGD	0.983
FIM	0.990	XEU	0.985
FRF	0.989	ZAR	0.992

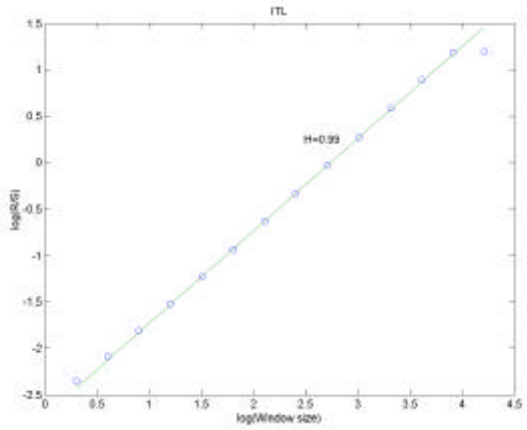

Figure 20. Estimated H, using the standard scaled window variance method of the Italian Lira-US dollar foreign exchange time series.

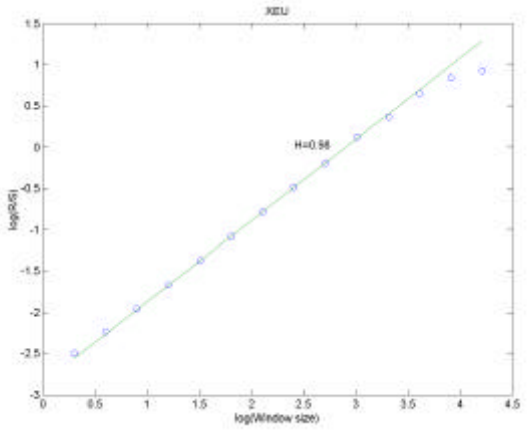

Figure 21. Estimated H, using the standard scaled window variance method of the Euro-US dollar foreign exchange time series.

4.1.2. Power Spectrum

The Fourier transform of a function *s(t)* is given by:

$$\tilde{s}(f) = \frac{1}{\sqrt{2p}} \int_{-\infty}^{+\infty} s(t) e^{2pift}\, dt \qquad (11)$$

and that of a finite, discrete time series by

$$\tilde{s}_k = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} s_j e^{2pikn/N} \qquad (12)$$

Here, the frequencies in physical units are $f_k = k/(NDt)$, where $k=-N/2,\ldots,N/2$ and Dt is the sampling interval (1/2 hour in our case). The power spectrum of a process is defined to be the squared modulus of the continuous Fourier transform, $P(f) = \left|\tilde{s}(f)\right|^2$. The power spectrum is particularly useful for studying the main frequencies in a system, since there will be sharper or broader peaks at the dominant frequencies and their integer multiples, the harmonics.

In Figures 22-25 we observe, as an example, the power spectrum of four of the foreign exchange time series. For all of them, we have found a behaviour of the type

$$P(f) \propto \frac{1}{f^a} \qquad (13)$$

where α is a positive real number. Table 3 summarises the values found for α.

Table 3. Estimated α for the foreign exchange time series, using the Welsch's method to calculate the power spectral density.

Data set	a	Data set	a
AUD	1.52	GBP	1.53
BEF	1.21	ITL	1.25
CAD	1.52	MYR	1.44
CHF	1.58	JPY	1.60
DEM	1.64	NLG	1.43
DKK	1.54	SEK	1.36
ESP	1.25	SGD	1.13
FIM	1.45	XEU	1.47
FRF	1.51	ZAR	1.55

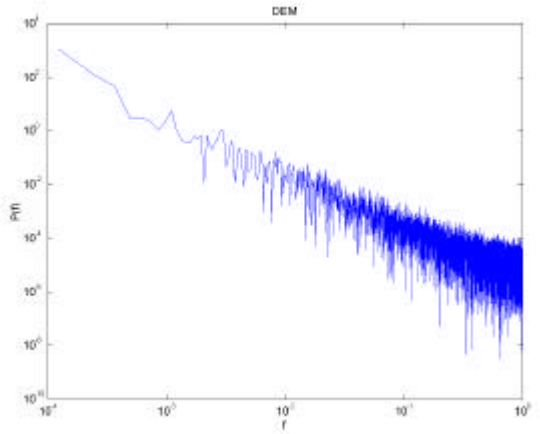

Figure 22. The power spectrum (log-log scale) of the German Mark-US dollar foreign exchange time series.

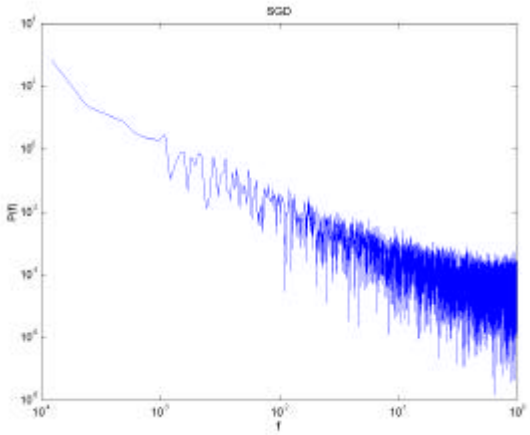

Figure 24. The power spectrum (log-log scale) of the Singapore dollar-US dollar foreign exchange time series.

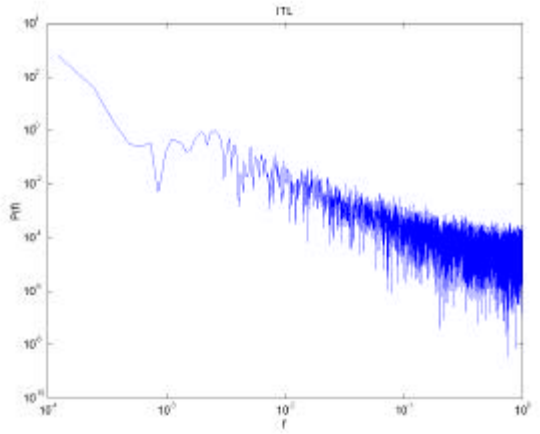

Figure 23. The power spectrum (log-log scale) of the Italian Lira-US dollar foreign exchange time series.

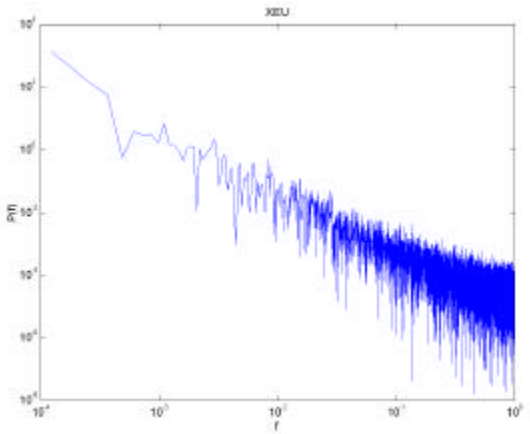

Figure 25. The power spectrum (log-log scale) of the Euro-US dollar foreign exchange time series.

It has been observed experimentally (Schuster, 1995) that the power spectra of a large variety of physical systems diverge at low frequencies with a power low $1/f^a$ ($0.8<\alpha<1.4$). This phenomenon is called 1/f-

noise. This allows a preliminary classification of the exchange time series in α<1.4, i.e. BEF, ESP, ITL, MYR, NLG, SEK, SGD and α>1.4, i.e. AUD, CAD, CHF, DEM, DKK, FIM, FRF, GPB, JPY, XEU, ZAR. In this case the ECU (XEU) is at the limit between both types with a value of α =1.47, and the maximum and minimum values are the German Mark (DEM) with α =1.64 and the Singapore dollar (SGD) with α =1.13, respectively. The appearance of a scaling behaviour in the power spectrum of economic time series support further, according to Theiler (1991), the existence of a self-organisation with many degree of freedom for these economic time series, but as we will see later on, since the time series are not stationary, this also can be interpreted as a sign of non-stationarity.

4.2. Finding the time delay and embedding dimension

Determining the time delay and the embedding dimension is considered as one of the most important steps in nonlinear time series modelling and prediction. A number of methods have been developed in determining the time delay and the minimum embedding dimension since the early beginning of nonlinear time series study. Here we will describe and apply several of them to the foreign exchange time series data sets.

4.2.1.Time delay

The first step in phase space reconstruction is to choose an optimum delay parameter t. Different prescriptions have appeared in the literature to choose t but they are all empirical in nature and do not necessarily provide appropriate estimates:

- *First passes through zero of the autocorrelation function*: In earlier works (Mees *et al*., 1987) it was suggested to use the value of t for which the autocorrelation function

$$C(\tau) = \sum_n [s(n) - \bar{s}][s(n+\tau) - \bar{s}] \qquad (14)$$

first passes through zero which is equivalent to requiring linear independence.

- *First minimum of the Average mutual information*: Fraser and Swinney (1986) suggested to use the average mutual information (AMI) function, *I(t)*, as a kind of nonlinear correlation function to determine when the values of *s(n)* and *s(n+ t)* are independent enough of each other to be useful as coordinates in a time delay vector but not so independent as to have no connection which each other at all. For a discrete time series, *I(t)* can be calculated as,

$$I(\tau) = \sum_{n,n+T} P(s(n), s(n+\tau)) \log_2 \left[\frac{P(s(n), s(n+\tau))}{P(s(n))P(s(n+\tau))} \right] \qquad (15)$$

where *P(s(n))* refers to individual probability and *P(s(n),s(n+ t))* is the joint probability density. Following the method developed by Abarbanel (1996), to determine *P(s(n))* we simply project the values taken form *s(n)* versus *n* back onto the *s(n)* axis and form an histogram of the values. Once normalised, this gives us *P(s(n))*. For the join distribution of *s(n)* and *s(n+ t)* we form the two-dimensional histogram in the same way.

In general, the time lag provided by *I(t)* is normally lower than the one calculated with the *C(t)*, $t_{AMI} \leq t_{correl}$, and provides the appropriate characteristic time scales for the motion. Even though *C(t)* is the optimum linear choice from the point of view of predictability in a least square sense of *s(n+ t)* from

knowledge of *s(n)*, it is not clear why it should work for nonlinear systems and it has been shown that in some cases it does not work at all.

4.2.2. Embedding dimension

The dimension, where a time delay reconstruction of the system phase space provides a necessary number of coordinates to unfold the dynamics from overlaps on itself caused by projection, is called the embedding dimension, d_E. This is a global dimension, which can be different from the real dimension. Furthermore, this dimension depends on the time series measurement, and hence, if we measure two different variables of the system, there is no guarantee that the d_E from time delay reconstruction will be the same from each of them.

The usual method for choosing the minimum embedding dimension is to compute some invariants of the attractor. By increasing the embedding dimension used for the computations, one notes when the value of the invariant stops changing. Since these invariants are geometric properties of the dynamics, they become independent of d for $d \quad d_E$, i.e. after the geometry is unfolded.

In this work, we have used two methods:

- *False Nearest Neighbours*: The method of False Nearest Neighbours (FNN) was developed by Kennel *et. al* (1992). In this case, the condition of no self-intersection states that if the dynamics is to be reconstructed successfully in R^d, then all the neighbour points in R^d should be also neighbours in R^{d+1}. The method checks the neighbours in successively higher embedding dimensions until it finds only a negligible number of false neighbours when increasing dimension from d to $d+1$. This d is chosen as the embedding dimension.

It was found by Kennel *et al.* (1992) that if the data set is clean from noise, the percentage of false nearest neighbours will drop from nearly 100% in dimension one to strictly zero when d_E is reached. Further, it will remain zero from then on since the dynamics is unfolded. If the signal is contaminated with noise (infinite dimension signal) we may not see the percentage of false nearest neighbours drop to near zero in any dimension. In this case, depending on the signal to noise ratio the determination of d_E will degrade.

- *E1 & E2 Method* : The method of FNN has some subjectivity in defining that a neighbour is false since the values of two threshold parameters have to be defined, Kennel *et. al* (1992). To improve this situation, Cao (1997) developed a similar method, which is based on evaluating the mean value of the distance between time-delay vectors, *E1(d).* However, if we look only at the quantity *E1(d)* we can obtain wrong results in the case of random signals. For time series data from a random set of numbers *E1(d),* in principle, will never reach a saturation value as *d* increases. But in practical computations, it is difficult to resolve whether *E1(d)* is slowly increasing or has stopped changing when d is sufficiently large. In fact, since available observed data samples are limited, it may happen that the *E1(d)* stop changing at some *d* although the time series is random. To solve this problem Cao (1997) suggested to consider the quantity *E2(d).* Let $y_i(d) = \{s(i), s(i+\tau), ..., s(i+(d-1)\tau\}$ and $y_{n(i,d)}$ the nearest neighbour of $y_i(d)$ in the d-dimensional reconstructed state space, then it is possible to define:

$$E^*(d) = \frac{1}{N - d\tau} \sum_{i=1}^{N-d\tau} \left| s_{i+d\tau} - s_{n(i,d)+d\tau} \right| \qquad (16)$$

$$E2(d) = \frac{E*(d+1)}{E*(d)} \qquad (17)$$

Since the future values are independent of the past values, *E2(d)*, for random data, will be equal to 1 for any *d*. However, for deterministic data, *E2(d)* is certainly related to *d*, and it cannot be a constant for all *d*. In other words, there must exist some *d*'s such that *E2(d)¹1*. The E1&E2 method depends only on the time delay, and the embedding dimension is calculated, as in the other methods, when the values of E1 and E2 reach saturation. Cao (1997) showed that the method does not strongly depend on how many points are available, provided there are enough and it can clearly distinguish between deterministic and stochastic.

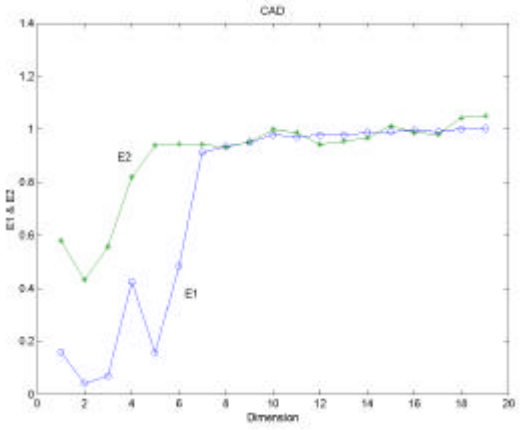

Figure 26. Embedding dimension calculation for the Canadian -US dollar foreign exchange time series.

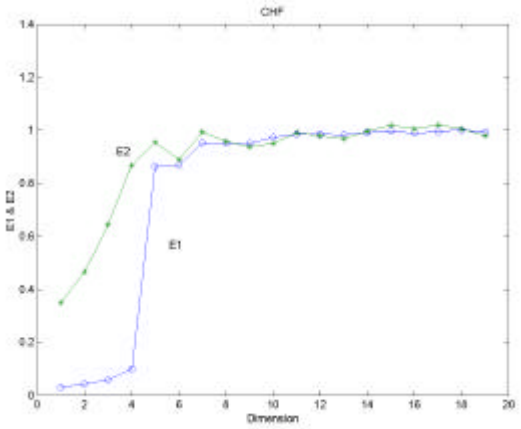

Figure 27. Embedding dimension calculation for the Swiss Franc -US dollar foreign exchange time series.

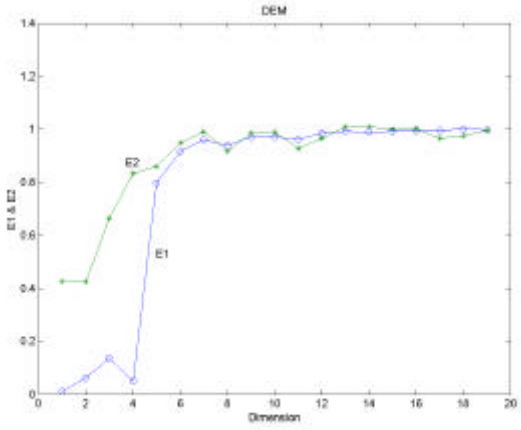

Figure 28. Embedding dimension calculation for the German Mark -US dollar foreign exchange time series.

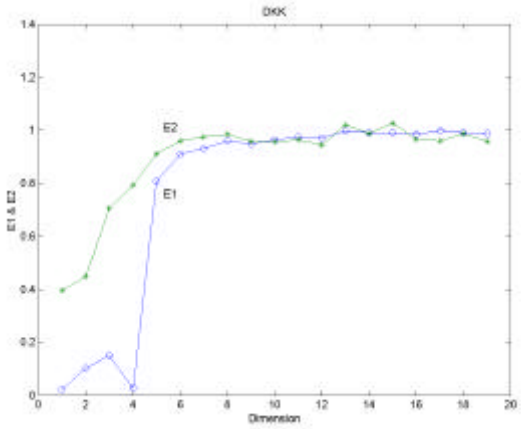

Figure 29. Embedding dimension calculation for the Danish Krone -US dollar foreign exchange time series.

Table 4. Time delay, t, and embedding dimension, d_E, found for the Foreign exchange time series data sets.

Data set	**t**	**d_E (FNN)**	**d_E (E1&E2)**	**Data set**	**t**	**d_E (FNN)**	**d_E (E1&E2)**
AUD	255	8	14	GPB	276	8	10
BEF	256	11	12	ITL	240	11	11
CAD	252	9	10	MYR	250	11	12
CHF	254	9	10	JPY	283	7	12
DEM	279	7	12	NLG	243	10	8
DKK	273	9	13	SEK	232	8	10
ESP	275	12	11	SGD	249	11	10
FIM	274	9	11	XEU	250	9	10
FRF	271	9	12	ZAR	270	15	14

Table 4 summarises the results obtained analysing the foreign exchange time series. The time delay have been obtained using the first minimum of the AMI, Eq. (15). The values found for the time delay correspond approximately to a week (5 days) of observations. This long time delay coincides with the time typical time delays obtained when analysing daily values, so from the dynamical systems point of view, using high frequency data means only that we are oversampling the time series. In addition, the embedding dimension have been computed using the methods of FNN (Kennel *et al*., 1992) and the E1&E2 method (Cao, 1997). The results of this last method are summarised in Figures 26-29, where it can be seen by looking at the value of *E2* that the time series analysed does not behave as stochastic signals, i.e. $E2 \approx 1$ for all d. Furthermore, all of them have high dimensionality, $d_E \geq 7$. This high values are in agreement with similar analysis carried out by Cao (2001) for other economic time series, i.e. daily variations in the British Pound and Japanese Yen/US dollar.

4.3. Detecting non-stationarity

As it has been mentioned before, broadly speaking a time series is said to be stationary if there is no systematic change in mean (no trend), in variance, and, if strictly periodic variations have been removed. Most of the probability theory of time series is concerned with stationary time series, and for this reason time series analysis often requires one to turn a non-stationary series into a stationary one so as to use this theory. However, it is also worth stressing that the nonstationary components, such as the trend, may sometimes be of more interest than the stationary residual.

We only report here a relatively simple stationarity test, called *space time separation plot*, introduced by Provenzale *et al*. (1992). The idea behind is that in the presence of temporal correlation the probability that a given pair of state points in the reconstructed state space, $\{s(t_i), s(t_i-Dt), s(t_i-2Dt),\ldots\}$, has a distance smaller than r, i.e. $\|\mathbf{s}_i-\mathbf{s}_j\| < r$, does not depend only on the position of the state but also on the time that has elapsed between them. This dependence can be detected by plotting the number of neighbour points as a function of two variables, the time separation and the spatial distance. In principle, one can create for each time separation an accumulated histogram of spatial distances. In the case of power-law noises the only points with small spatial separation are dynamically near neighbours, i.e. the series is non-recurrent in phase space. In this case the countour curves do not saturate. In the case of stationarity, we will find a saturation in the plot.

Figures 30-33 show the results of the test for some of the analysed time series. In those graphics the separation time is represented in the horizontal axis whereas the base 2 logarithm of the separation in space is

represented in the vertical axis. For small *Dt* points are always near neighbours in space, as their time separation increases so does their separation in space (Provenzale *et al*. 1992). Technically we have to create, for each time separation *Dt* an accumulated histogram of spatial distance *e*. We have used the program *stp* of Tisean (Kantz and Schreiber, 1997) which returns level lines for 10%, 20%, …of the pairs with a given temporal separation *Dt*.

As it can be seen the curves do not saturate at all. Apart from the non-stationarity, this is another indication that the data we are analysing have significant power in the low frequency, such as 1/f noise or Brownian motion. In this case, all points in the data set are temporally correlated and there is no way of determinig an attractor dimension from the sample. A similar situation arises if the data set is too short. Then there are no pairs left after removing those which are temporally correlated. If we regard the problem from a different point of view, correlation times of the order of the length of the sample (nonsaturating curves) means that the data do not sample the observed phenomenon sufficiently (Kantz and Schreiber, 1997).

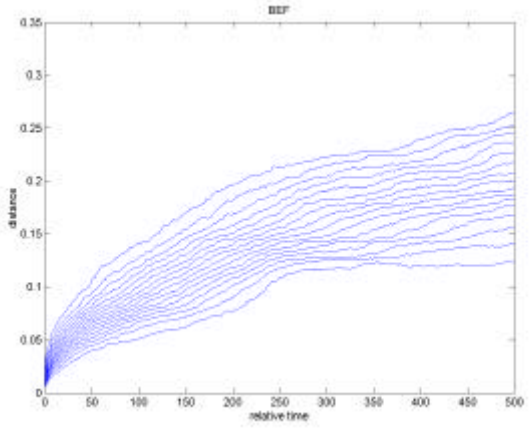

Figure 30. Space-time separation plot Belgium Franc-US dollar foreign exchange time series.

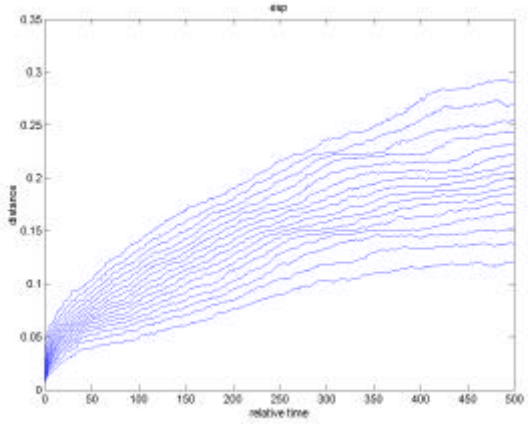

Figure 31. Space-time separation plot of Spanish Peseta-US dollar foreign exchange time series.

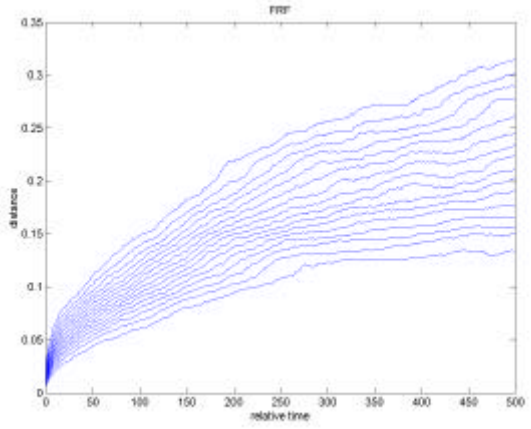

Figure 32. Space-time separation plot of French Franc-US dollar foreign exchange time series.

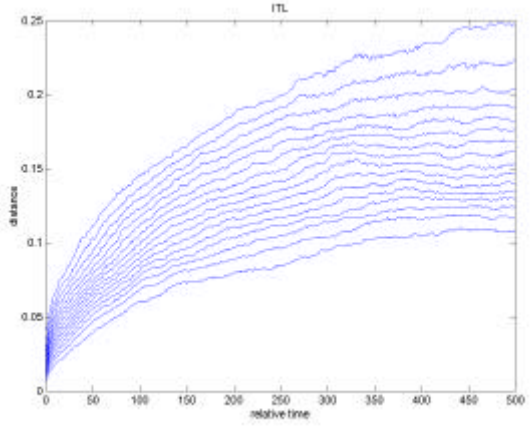

Figure 33. Space-time separation plot of Italian Lira-US dollar foreign exchange time series.

4.4. Testing for non-linearity (Surrogate data test)

If the dynamics that has generated the time series is not known or if the data are noisy, it is important to investigate whether the amount of nonlinear deterministic dependencies is worth analyzing further or whether the series can be considered as stochastic. Hence, one of the first steps before applying nonlinear techniques to high frequency stock exchange data is to investigate if the use of such advanced techniques is justified by the data. The main reason behind this reasoning is that linear stochastic processes can create very complicated looking signals and that not all the structures that we find in a data set are likely to be due to nonlinear dynamics going on within the system. The method of surrogate data, see for example Schreiber and Schmitz (2000) for a review, has become a useful tool to address the question if the irregularity of the data is most likely due to nonlinear deterministic structure or rather due to random inputs to the system or fluctuations in the parameters.

The method of surrogate data, which was first introduced by Theiler *et al*. (1992) in nonlinear time series analysis, consists of generating an ensemble of "surrogate" data sets similar to the original time series, but consistent with the null hypothesis, usually that the data have been created by a stationary Gaussian linear process, and of computing a discriminating statistic - normally based on the forecasting error - for the original and for each of the surrogate data sets.

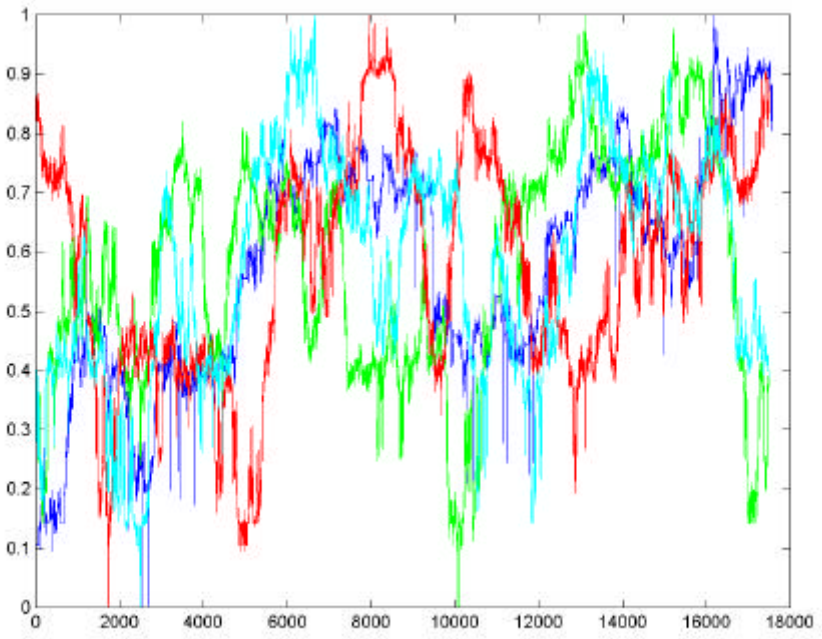

Figure 34. Original time series (BEF) in blue and three surrogate time series generated using the TISEAN software package (Hegger *et al.*, 1999) with *surrogates* program.

In order to test the null hypothesis with a 95% significance level, we have created 19 surrogate data sets for each foreign exchange time series, which are comparable to the measured data in the sense that the surrogate data have the same mean, variance and Fourier power spectrum (Schreiber and Schmitz, 1996), but the data

comes from a stationary linear stochastic process with Gaussian inputs, see fig. 34 for an example. The statistics we have used to reject the null hypothesis is the error in the nonlinear prediction.

Table 5. Results form the surrogate data test. Rejection of the null hypothesis means that the non-linear prediction error found in the original time series is lower than in the surrogate time series generated by a stationary Gaussian linear process.

Data set	**Surrogates test**	**Data set**	**Surrogates test**	**Data set**	**Surrogates test**
AUD	yes	ESP	yes	JPY	yes
BEF	yes	FIM	no	NLG	yes
CAD	no	FRF	yes	SEK	no
CHF	yes	GPB	no	SGD	yes
DEM	no	ITL	no	XEU	yes
DKK	no	MYR	no	ZAR	yes

We have used the program *surrogates* within TISEAN program (Hegger *et al*., 1999) to generate the 19 surrogate data sets and the program *predict* to compute the prediction error. Table 5 summarizes the results. As it can be seen, 10 data sets from 18 reject the null hypothesis, which means that in all these cases the nonlinear zero-order prediction error was lower in the original time series that in the surrogate data set. These results are in agreement with the previous findings of the space time separation plot in which one can see that the curves does not saturate which means that the system is either stochastic or it is not stationary. If we are testing using an inappropriate null hypothesis it is clear that we are going to obtain mixed results. In some cases it is possible (if one knows the type of non-stationarity) to generate a null hypothesis that takes this into account but unfortunately this is not the case with foreign exchange time series.

4.5. Recurrence quantification analysis (RQA)

The actual methods developed in non-linear time series analysis assume that the data series under analysis have reach their attractors and there are not in a transient phase, that they are autonomous and their lengths are much longer than the characteristic time of the system in question. In the case of foreign exchange time series this does not happen and it may be useful to have another procedure to analyse these data.

Eckmann *et al*. (1987) introduced a new graphical tool, which they called a recurrence plot (RP). The recurrence plot is based on the computation of the distance matrix between the reconstructed points in the phase space, i.e. $\mathbf{s}_i=\{s(t), s(t-t), s(t-2t),...s(t+(d_E-1)t\}$,

$$d_{ij} = \left\| \mathbf{s}_i - \mathbf{s}_j \right\| \qquad (18)$$

This produces an array of distances in a NxN matrix, **D**, N being the number of points under study. Once this distance matrix is calculated, in the original paper of Eckmann *et al*. (1987), it was displayed by darkening the pixel located at specific (i,j) coordinates which correspond to a distance value between i and j lower than a predetermined cutoff, i.e. a ball of radius r_i centered at $\mathbf{s}_i$. Requiring $r_i = r_j$, the plot is symmetric and with a darkened main diagonal correspondent to the identity line. The darkened points individuate the recurrences of the dynamical system and the recurrent plot provides insight into periodic structures and clustering properties that are not apparent in the original time series (Eckmann *et al*.,1987).

In order to extend the original concept and made it more quantitative Zbilut and Webber (1992) developed a methodology called Recurrence Quantification Analysis (RQA) (Webber and Zbilut, 1994). As a result, they

defined several variables to quantify RPs, namely: *%recur* (percentage of darkened pixels in recurrence plot); *%deter* (percentage of recurrent points forming diagonal line structures); *entropy* (Shannon entropy of line segments distributions); *trend* (measure of the paling recurrence points away from the central diagonal); $1/line_{max}$ (reciprocal of the longest diagonal line segment which relates directly to largest positive Lyapunov exponent) (Trulla *et al*., 1996). These five recurrence variables quantify the deterministic structure and complexity of the plot: *%recur* quantifies the amount of cyclic behaviour; *%deter* the amount of determinism through the counting of "sojourn points" (Gao and Zheng, 1994); *entropy* the richness of deterministic structuring; $1/line_{max}$ scales with the maximum Lyapunov exponent; while *trend* is essentially a measure of nonstationarity.

Since RQA methodology is, in principle, independent of limiting constraints such as data set size, data stationarity, and of assumptions regarding statistical distributions of data, RQA has been applied to physiological systems characterized by non-homeostatic transient and state changes. Furthermore, it has also be extended the application areas ranging from analysis of molecular dynamics simulation data (Manetti *et al*.,1999) to human speech (Orsucci *et al*., 1999).

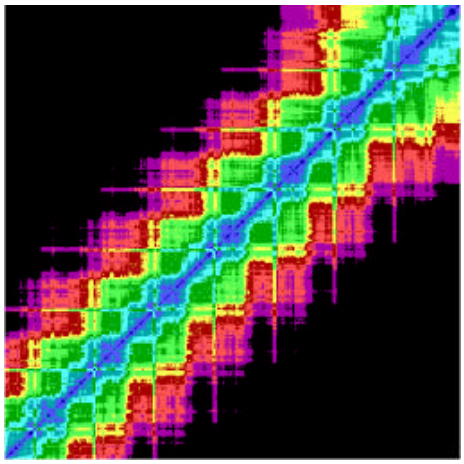

Figure 35. RP Analysis of the Australian -US dollars foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

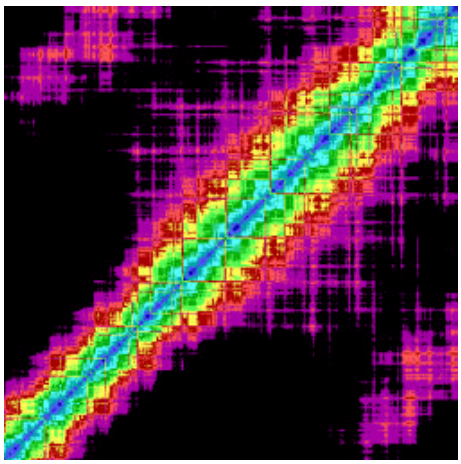

Figure 36. RP Analysis of the Belgium Franc-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

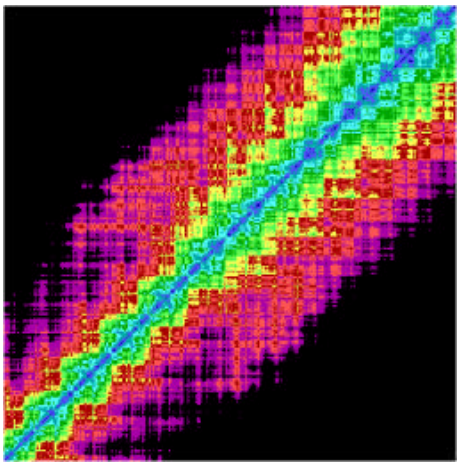

Figure 37. RP Analysis of the Canadian-US dollars foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day)..

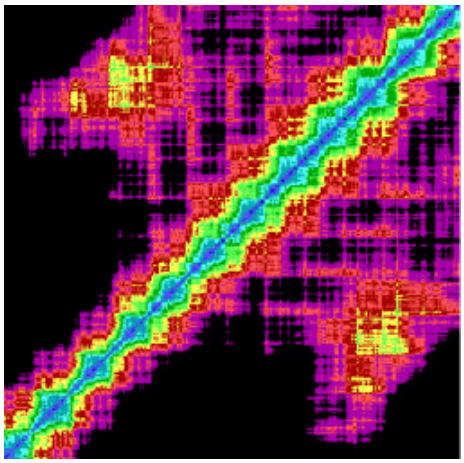

Figure 38. RP Analysis of the Swiss Franc-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

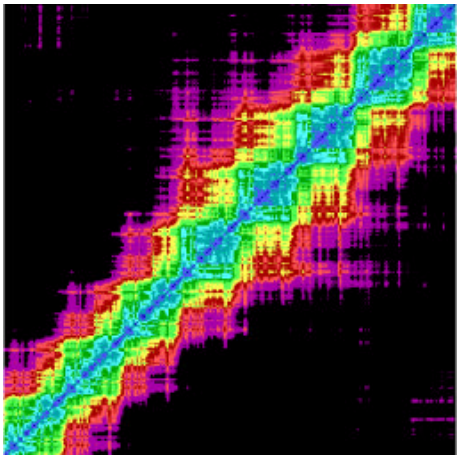

Figure 39. RP Analysis of the German Mark-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

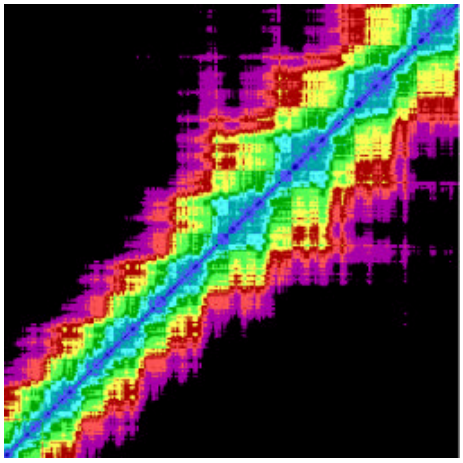

Figure 40. RP Analysis of the Danish Krone-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

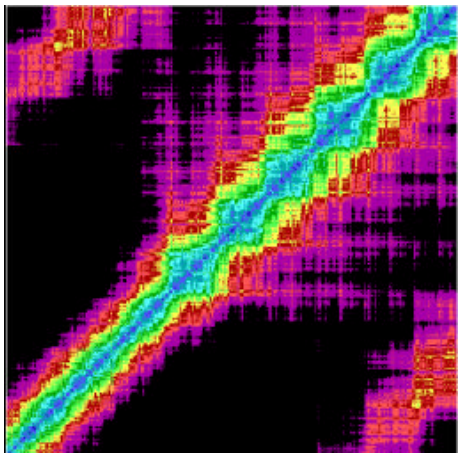

Figure 41. RP Analysis of the Spanish Peseta-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

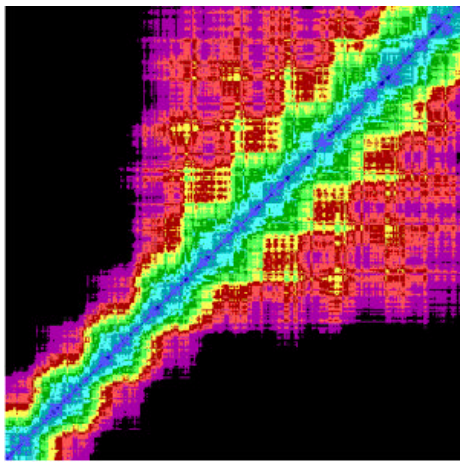

Figure 42. RP Analysis of the Finnish Markka-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

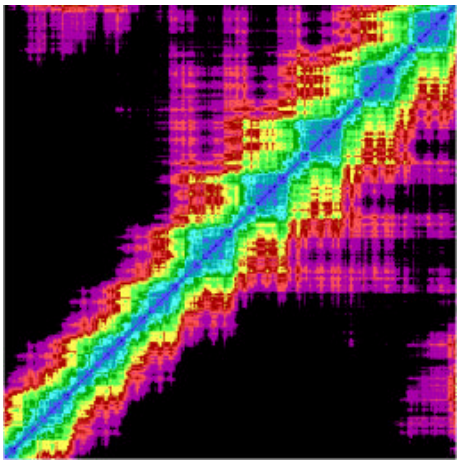

Figure 43. RP Analysis of the French Franc-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

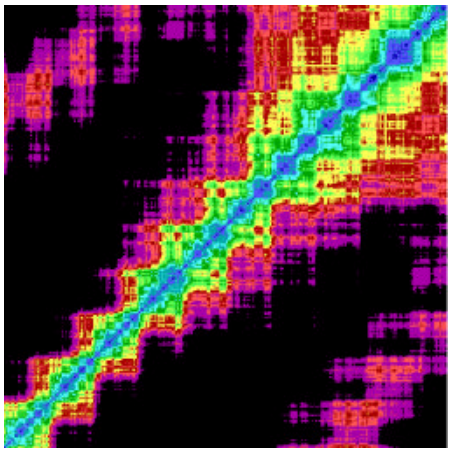

Figure 44. RP Analysis of the British Pound-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

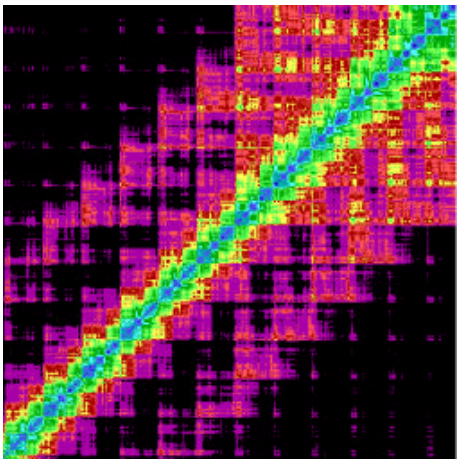

Figure 45. RP Analysis of the Italian Lira-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

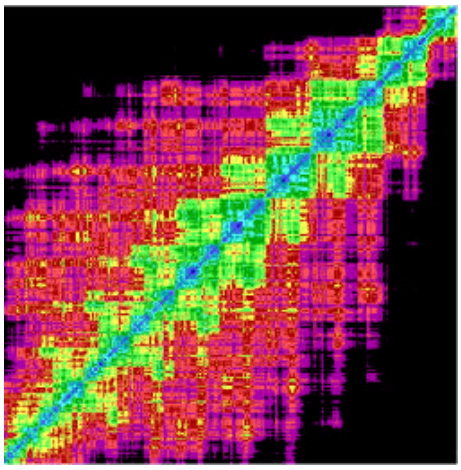

Figure 46. RP Analysis of the Malaysian Ringgit-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

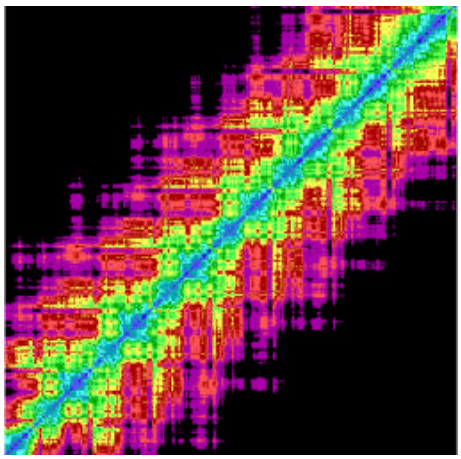

Figure 47. RP Analysis of the Japanese Yen-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

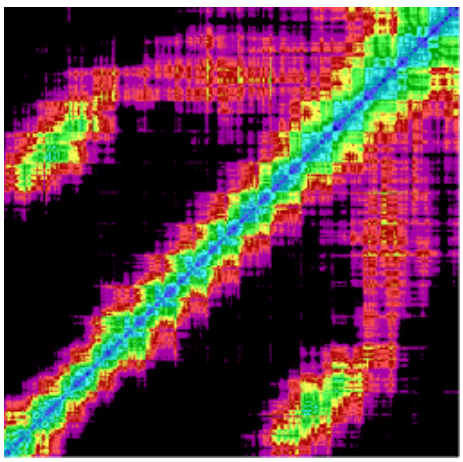

Figure 48. RP Analysis of the Duch Guilder-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

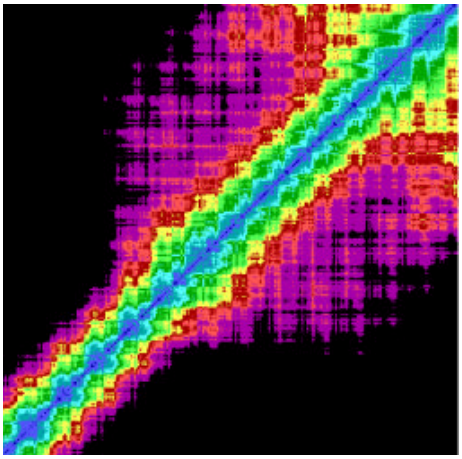

Figure 49. RP Analysis of the Swedish Krona-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

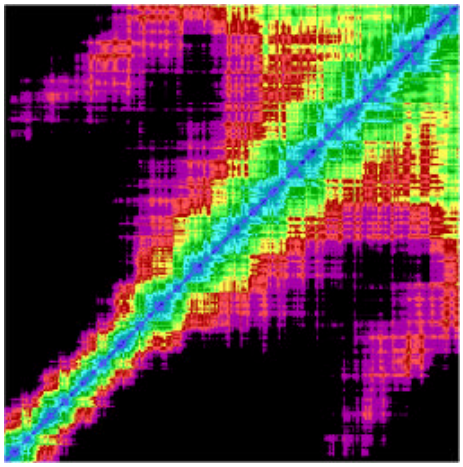

Figure 51. RP Analysis of the Euro-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

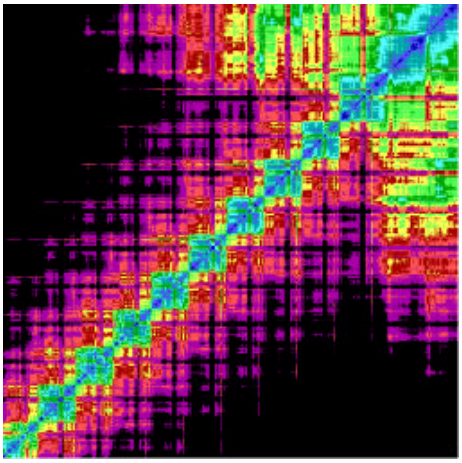

Figure 50. RP Analysis of the Singapore-US dollars foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

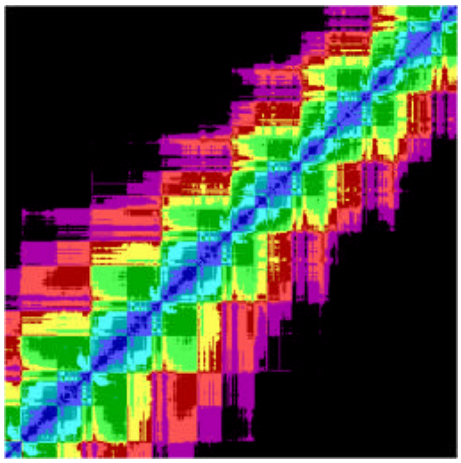

Figure 52. RP Analysis of the South African Rand-US dollar foreign exchange time series. RQA parameters: time delay and embedding dimension (Table 4), distance cutoff: mean distance between points, line definition: 48 points (1 day).

First RQA was performed on all the time series using the time delay and embedding dimension (E1&E2) from Table 4 for the state space reconstruction. The recurrence plots obtained are shown in Figs. 35-52, whereas the RQA computed values are summarised in Table 6. By visual inspection, it is possible to distinguish between several general patterns in the recurrence plots. This allows an empirical classification of the series as follows:

- Constant RP over time: AUD and JPY
- Higher recurrence during the first 6 months: MYR and ZAR
- Higher recurrence during the last 6 months: CAD, DEM, DKK, FIM, SEK,SGD
- Aeroplane structure: BEF, CHF, ESP, FRF, GPB, ITL, NLG, XEU

Of course, for some time series this classification is clearer than for others and there are some cases that fall in between.

Table 6. RQA values for the foreign exchange time series.

Data set	***%recur***	***%deter***	**entropy**	**maxline**	**trend**	**Data set**	**%recur**	**%deter**	**entropy**	**maxline**	**trend**
AUD	50.2	93.6	8.5	1943	-74.9	GBP	45.4	85.7	8.0	2774	-20.9
BEF	42.3	70.5	5.8	2552	-26.5	ITL	47.1	74.8	7.4	2858	-42.8
CAD	50.7	86.9	7.7	2990	-47.9	MYR	60.5	82.7	7.5	2508	-57.3
CHF	51.5	81.3	7.9	2972	-24.2	JPY	51.3	87.8	7.6	2145	-68.0
DEM	41.5	88.2	8.0	2189	-56.4	NLG	48.7	85.2	8.0	3557	-19.4
DKK	43.2	90.1	7.8	1982	-65.4	SEK	48.1	86.4	8.0	3170	-40.9
ESP	45.9	73.5	7.9	2508	-4.4	SGD	48.2	79.0	7.7	3017	-37.0
FIM	56.3	93.0	8.7	2480	-53.8	XEU	51.7	87.9	8.3	3008	-29.2
FRF	47.2	82.9	7.7	2227	-37.9	ZAR	52.3	90.7	6.8	1748	-84.0

Afterwards, RQA was repeatedly performed on 336-point epochs –which corresponds to a week of data-, generating 305 values for each of the five RQA variables, i.e. *%recur*, *%deter*, *entropy*, *trend*, $1/line_{max}$. Neighbouring epochs were shifted by 48 points, which corresponds to one day of data. The nonlinear variables *%recur*, *%deter* and $line_{max}$ are plotted in Figs. 53-70 as function of time. *%recur* quantifies the percentage of the plot occupied by recurrent points, which corresponds to the proportion of recurrent pairs over all possible pairs below the chosen radius. *%deter* denotes the percentage of recurrent points that appear in sequence, forming diagonal line structures in the distance matrix. It corresponds to the amount of patches of recurrent behaviour in the studied time series, which indicates portions of the state space where the system resides for a longer time than expected by chance alone (Zak *et al*., 1997). This is an important indication of deterministic signature in the time series. Following its variation it is possible to distinguish several levels and variations, and hence, it is possible to study how our forecasting capabilities would change with the level of *%deter* we observe in the time series. For example, there is a generalised increase in the *%deter* in summer around July (*time* = 10000) for a considerable number of exchange rates time series, mainly in the Euro zone. As underlined in the Historical background Section it was just in the summer of1996 that the EU member states decided to reduce their budgets deficits in order to meet the EMU criteria. $line_{max}$ is the length in terms of consecutive points of the longest recurrent line in the plot scales with the maximum Lyapunov exponent Trulla *et al*. (1996) found that. $line_{max}$ was able to accurately predict the value of the maximum Lyapunov exponent and locate bifurcation points in a logistic map going from regular to chaotic regime. For example, it is possible to see in Figs. 88, 91 and 97, that the entrance of the Finnish Markka on October 1996 (around 14000) and of the Italian Lira on November 1996 (around 16000), in the Exchange-Rate Mechanism changed $line_{max}$ not only in both currencies but also in the Euro time series.

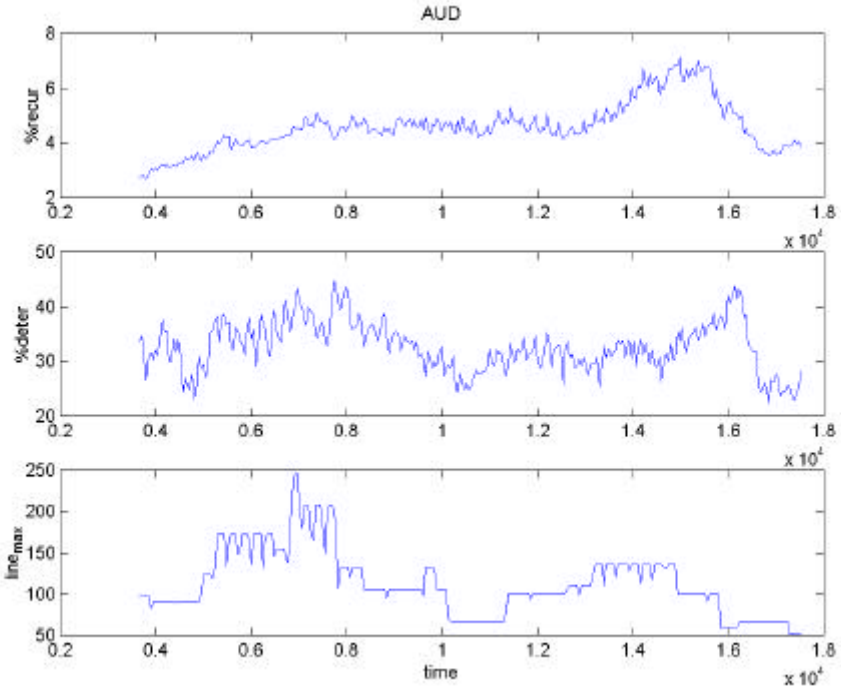

Figure 53. Nonlinear metrics of the Australian-US dollars foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

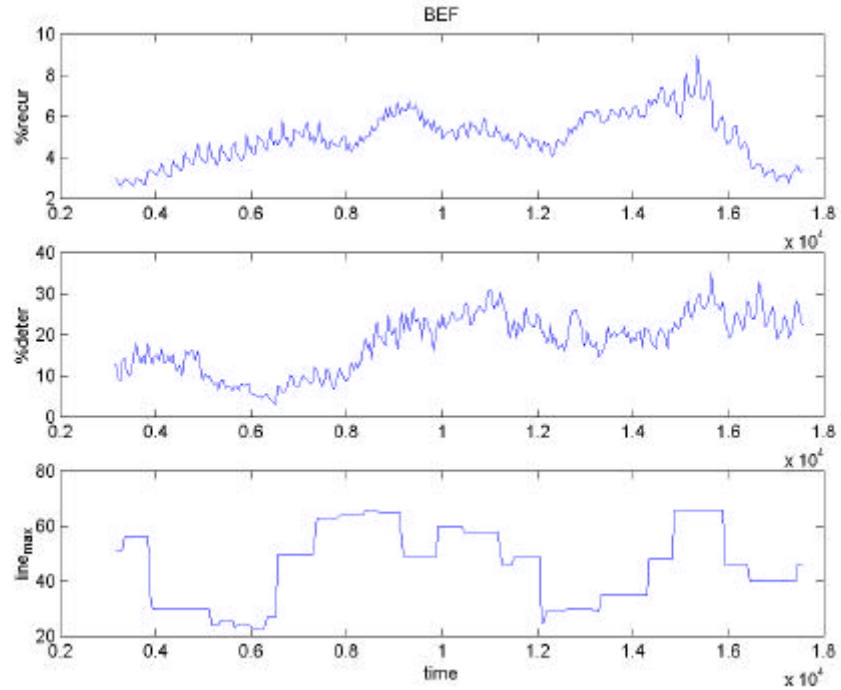

Figure 54. Nonlinear metrics of the Belgium Franc -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

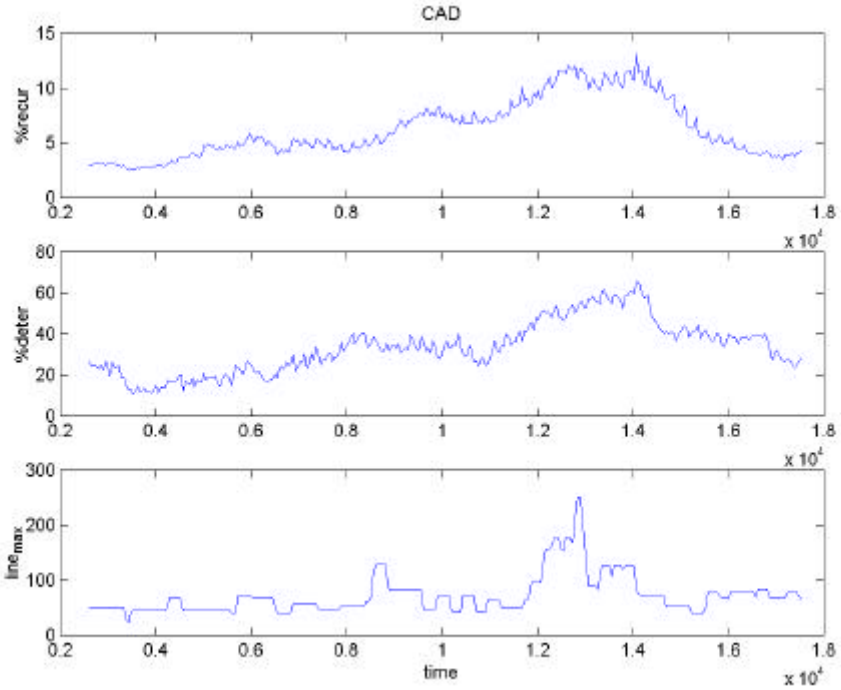

Figure 55. Nonlinear metrics of the Canadian-US dollars foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

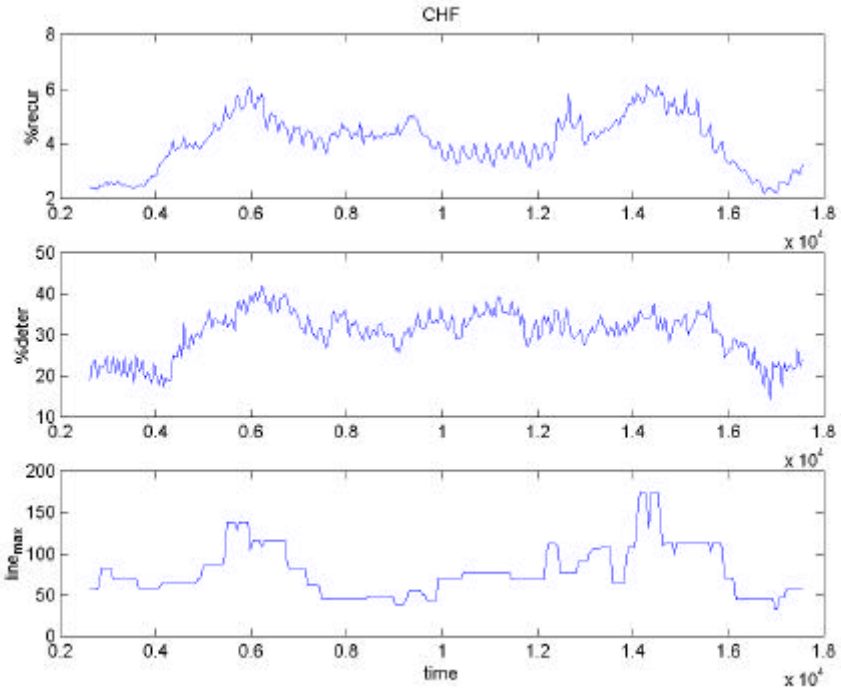

Figure 56. Nonlinear metrics of the Swiss Franc-US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

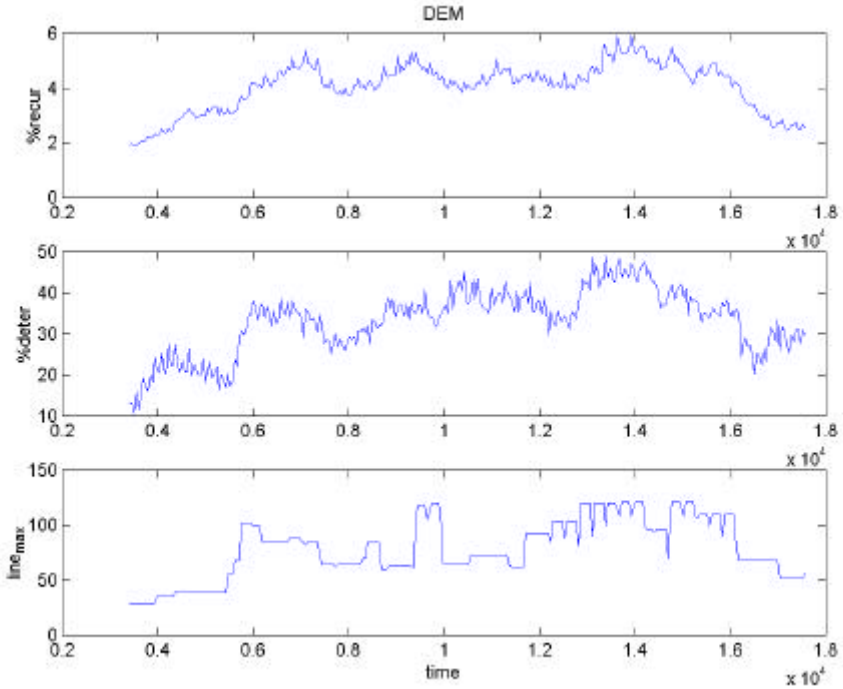

Figure 57. Nonlinear metrics of the German Mark -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

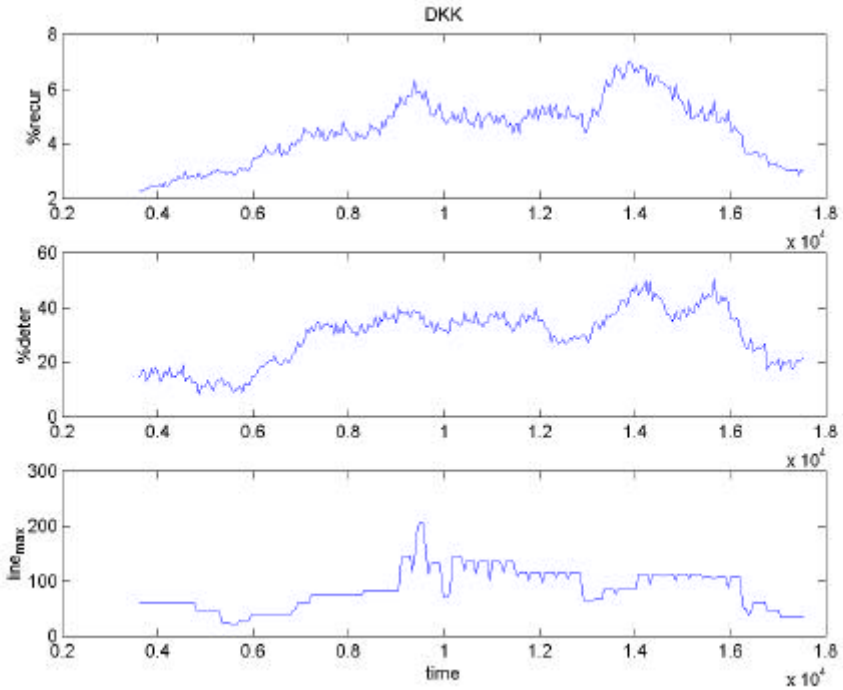

Figure 58. Nonlinear metrics of the Danish Krone -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

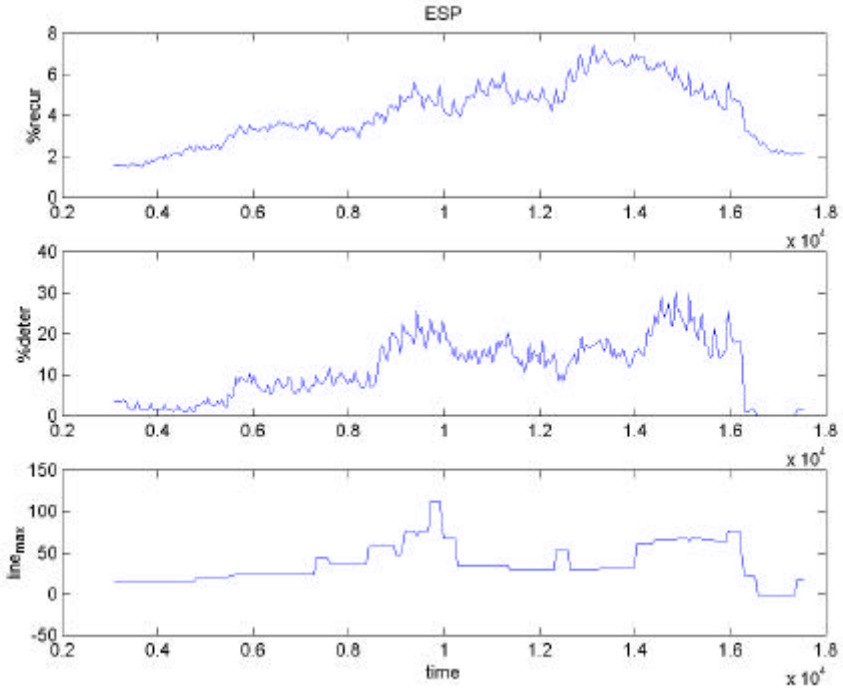

Figure 59. Nonlinear metrics of the Spanish peseta-US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

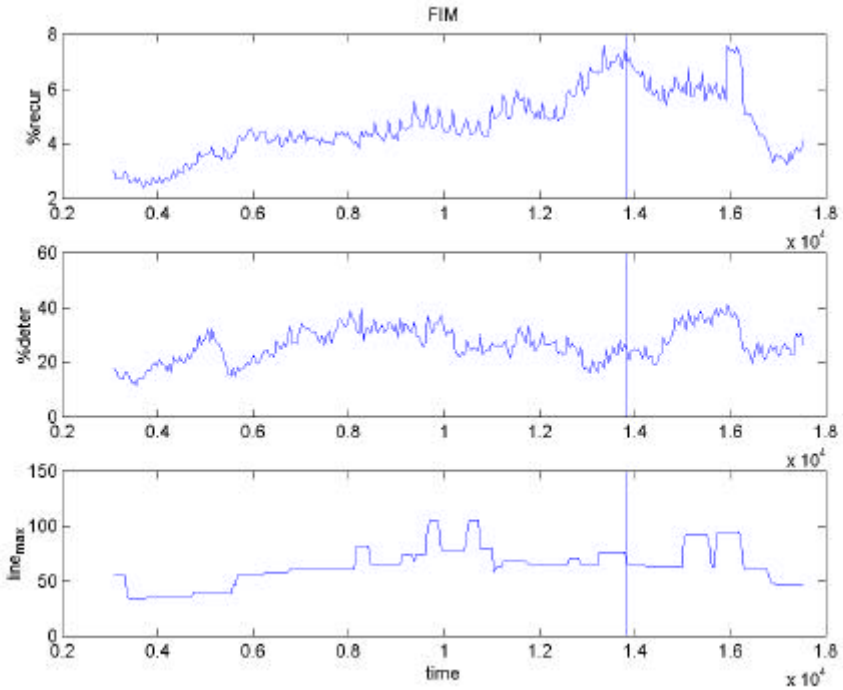

Figure 60. Nonlinear metrics of the Finnish Markka -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours). The vertical line represents the entry of the Finnish Markka in the European Monetary System.

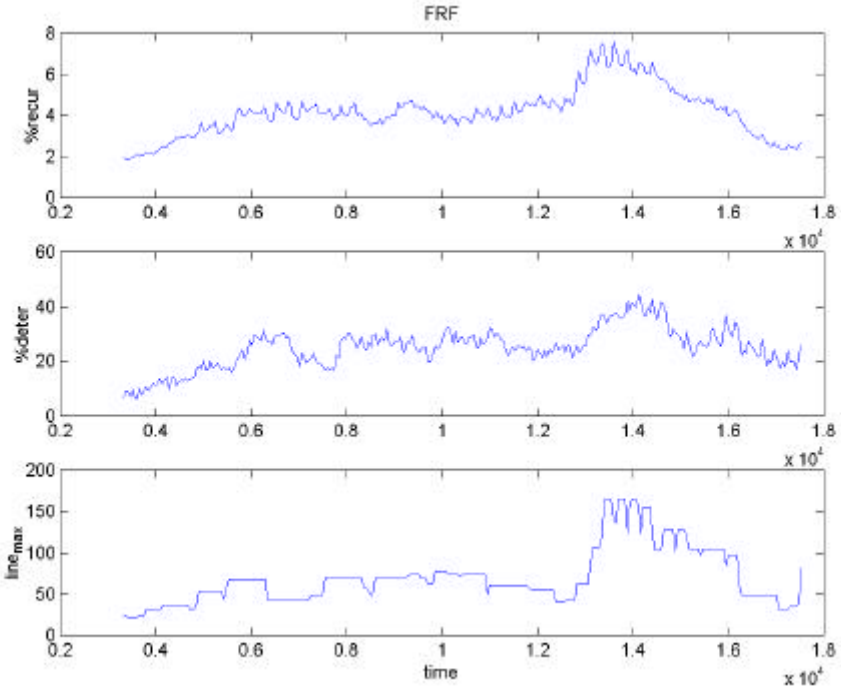

Figure 61. Nonlinear metrics of the French Franc -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

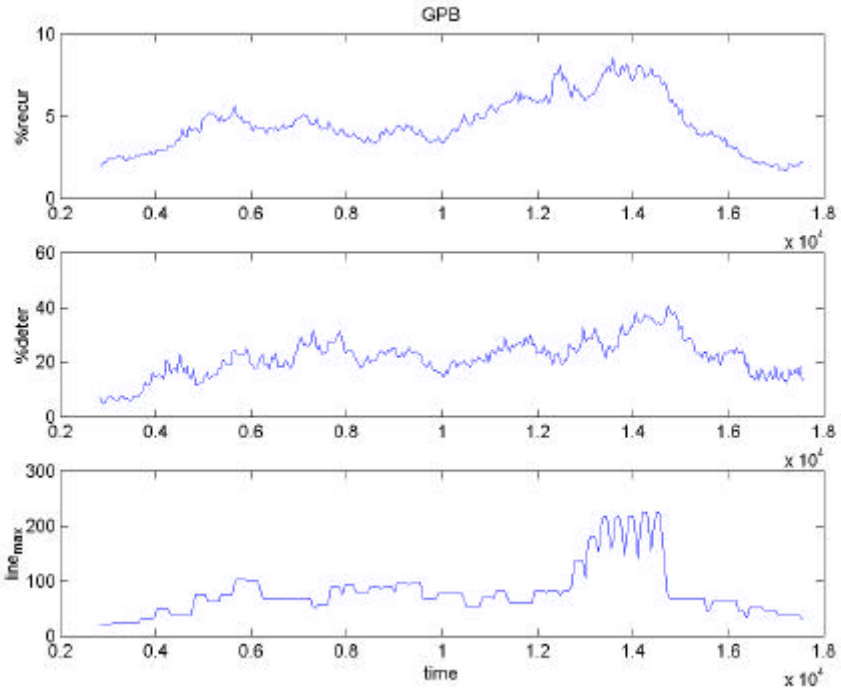

Figure 62. Nonlinear metrics of the British Pound -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

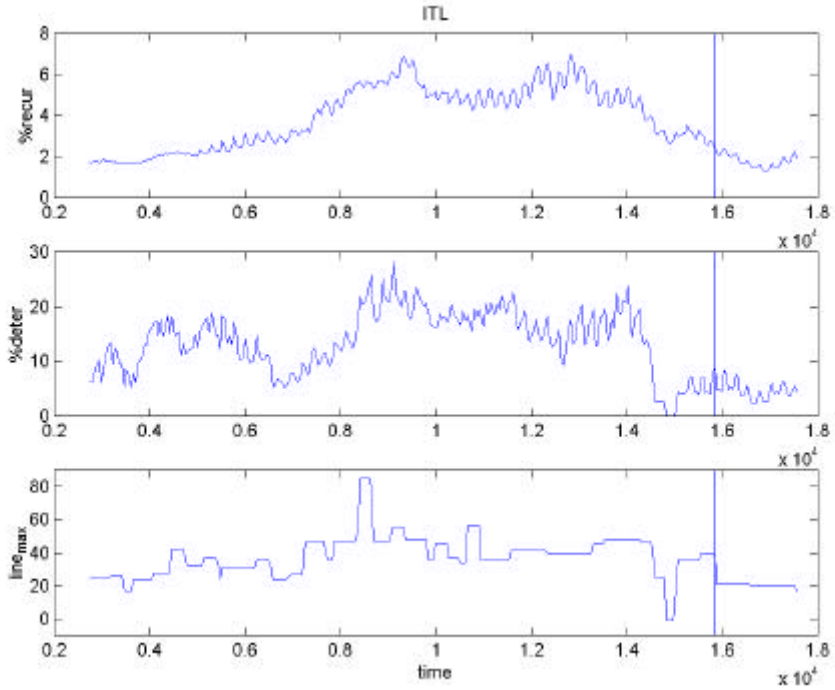

Figure 63. Nonlinear metrics of the Italian Lira -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours). The vertical line represents the entry of the Italian Lira in the European Monetary System.

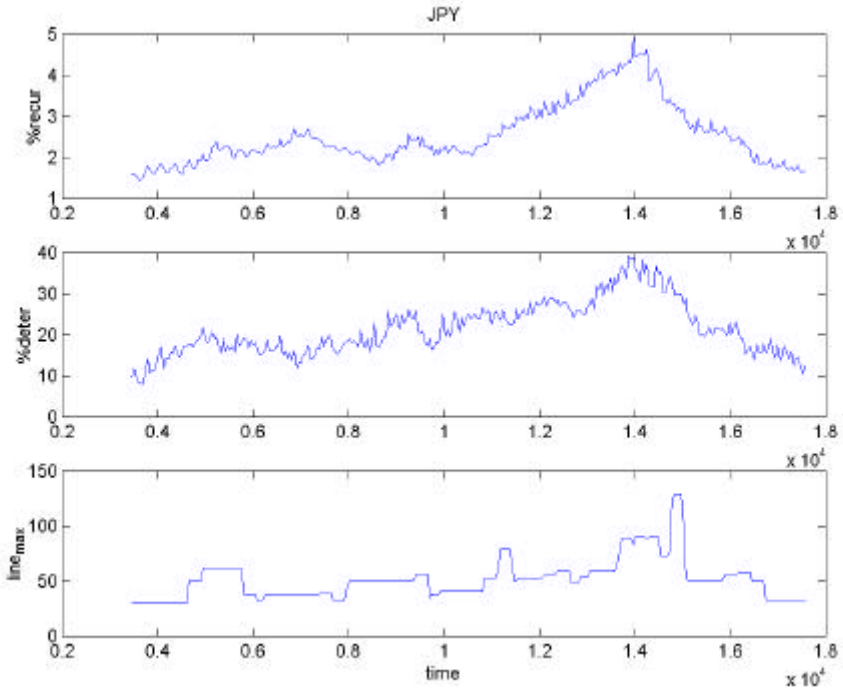

Figure 64. Nonlinear metrics of the Japanese Yen -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

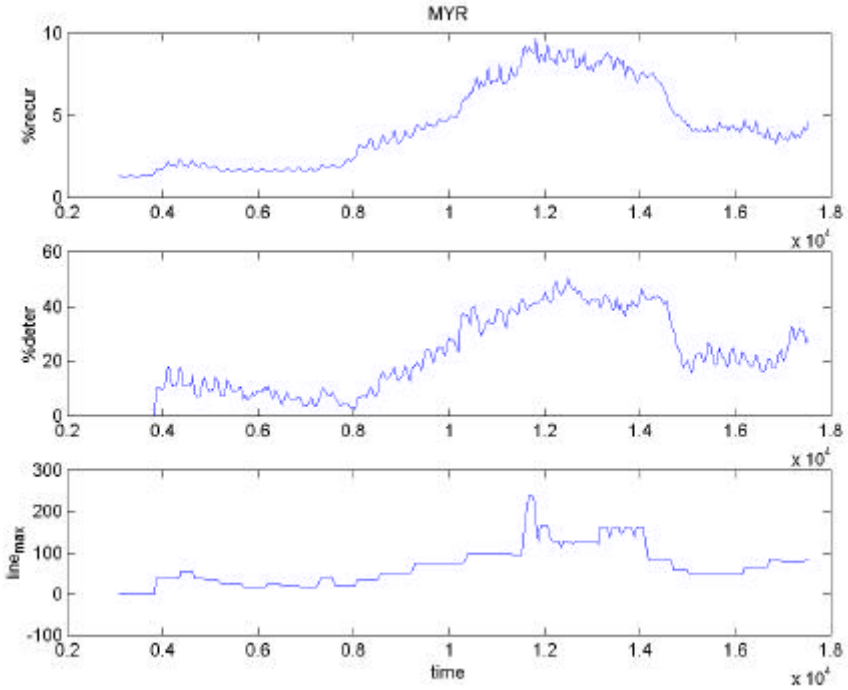

Figure 65. Nonlinear metrics of the Malaysian Ringgit -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

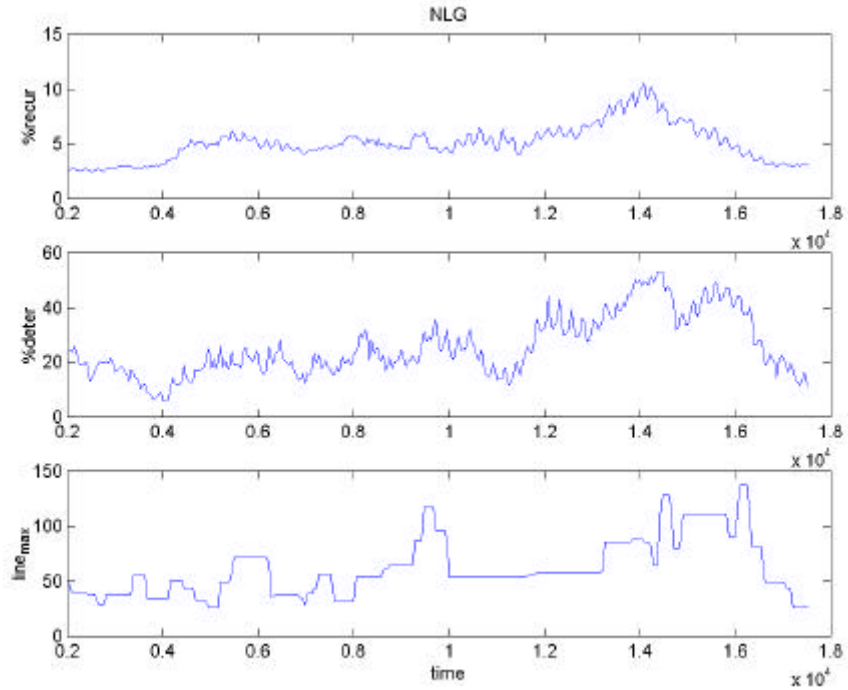

Figure 66. Nonlinear metrics of the Duch Guilder -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

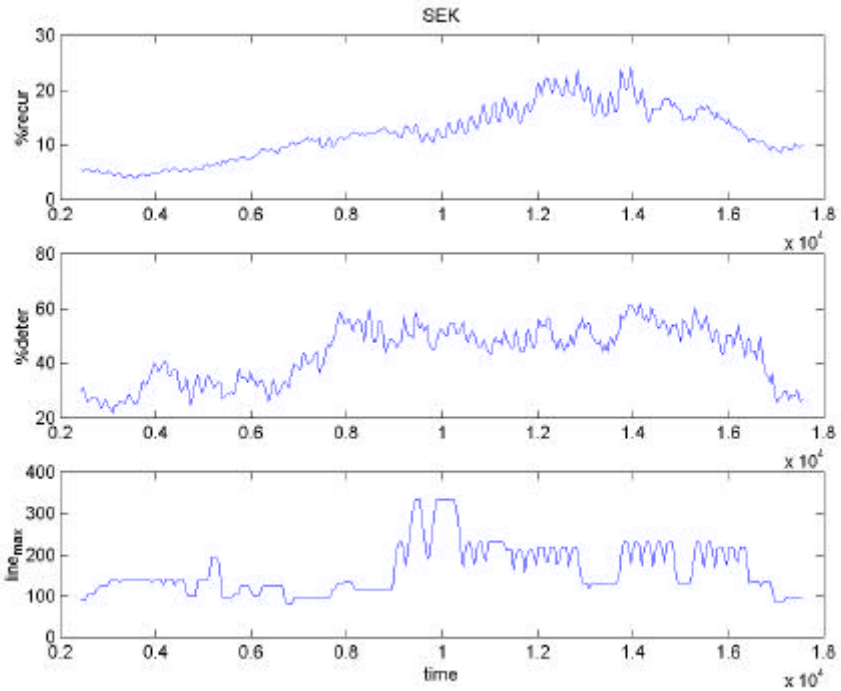

Figure 67. Nonlinear metrics of the Swedish Krona -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

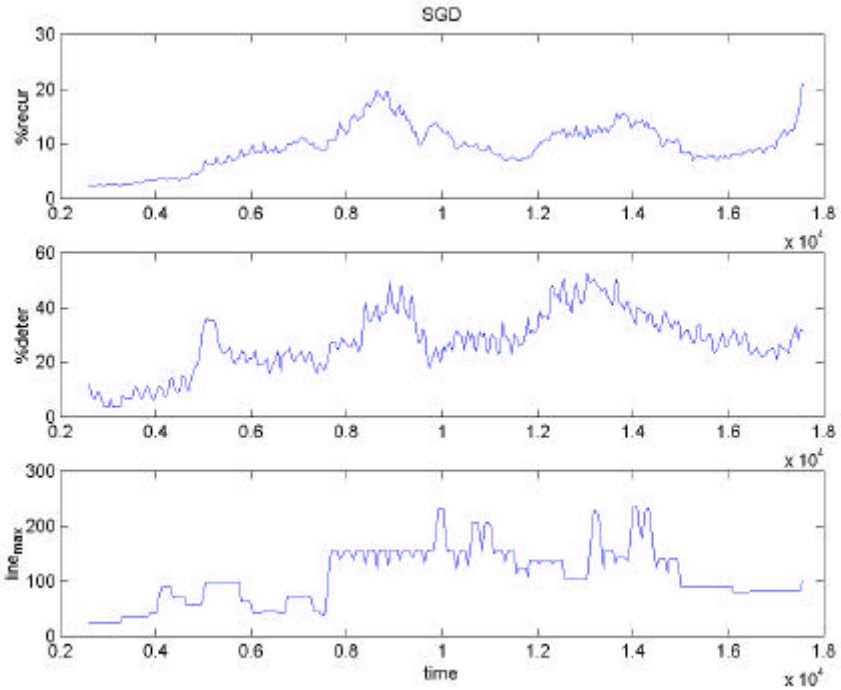

Figure 68. Nonlinear metrics of the Sigapore -US dollars foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

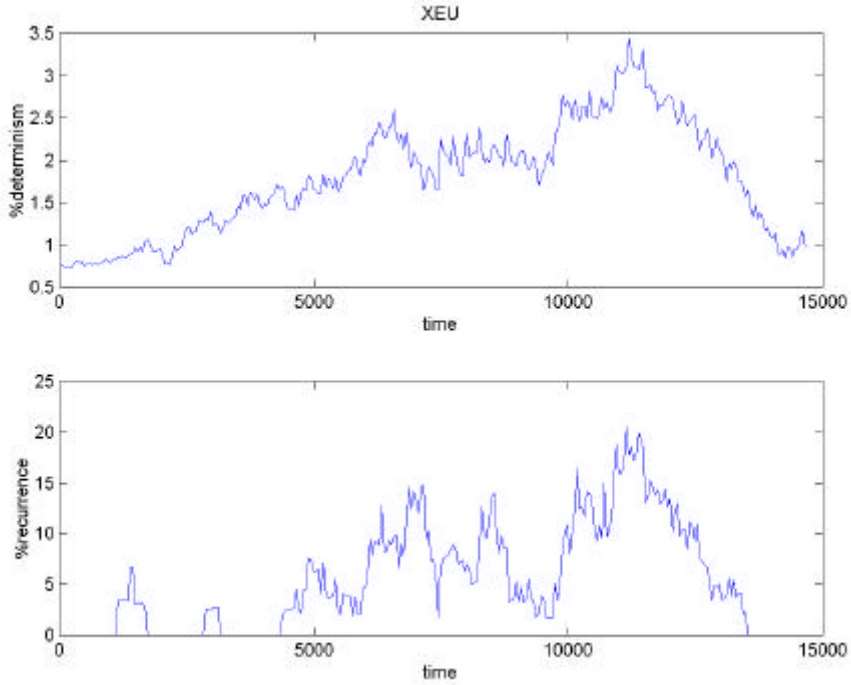

Figure 69. Nonlinear metrics of the Euro -US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours). The vertical lines represent the entry of the Finnish Markka and the Italian Lira in the European Monetary System, respectively.

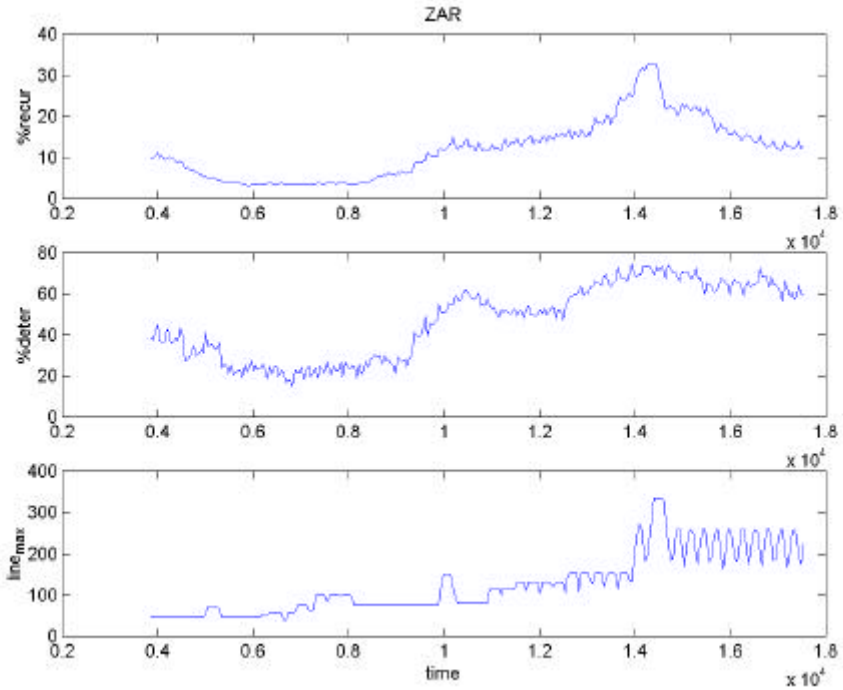

Figure 70. Nonlinear metrics of the South African Rand-US dollar foreign exchange time series: *%recur* (upper panel), *%deter* (middle panel) and $line_{max}$ (lower panel). Values are computed from a 336 point window (epoch), data are shifted 48 points between epochs. RQA parameters: time delay and embedding dimension (see Table 4), distance cutoff: mean distance between points/10, line definition: 16 points (8 hours).

5. Comparison between currency echange rates time series

In order to compare the different currency exchange rates time series, the *%recurrence* values were calculated using the same embedding parameters for the eighteen time series. A mean value for the time delay, t = 260, and embedding dimension, d_E =11, were used. The radius cutoff was defined as 10% of the mean distance between points in all the time series, (i.e. 0.043). Then the calculated *%recurrence* values were plotted against each other, this guarantees that we are comparing the same time periods in the currency exchange rates time series. The data sets were fitted by linear regression and the values of the correlation coefficients, *r*, are summarised in Table 7.

Even though Table 7 measures a linear correlation between *%recur* of the currencies, we speak about nonlinear correlation between exchange rate because the relationship between the time series and the *%recur* is nonlinear.

Table 7.Regression coefficients for the high frequency currency exchange rates time series.

	AUD	BEF	CAD	CHF	DEM	DKK	ESP	FIM	FRF	GPB	ITL	JPY	MYR	NLG	SEK	SGD	XEU	ZAR
AUD		0.835	0.602	0.615	0.840	0.798	0.840	0.817	0.859	0.512	0.482	0.686	0.488	0.879	0.731	0.508	0.782	0.533
BEF			0.689	0.465	0.936	0.953	0.890	0.810	0.924	0.554	0.751	0.667	0.604	0.927	0.815	0.719	0.930	0.596
CAD				0.325	0.726	0.788	0.767	0.633	0.750	0.841	0.726	0.905	0.801	0.697	0.751	0.560	0.818	0.868
CHF					0.584	0.405	0.441	0.363	0.590	0.562	0.158	0.551	0.010	0.554	0.144	0.001	0.367	0.163
DEM						0.945	0.859	0.834	0.935	0.654	0.742	0.728	0.617	0.903	0.756	0.686	0.906	0.574
DKK							0.899	0.818	0.931	0.642	0.842	0.725	0.731	0.885	0.856	0.754	0.965	0.683
ESP								0.865	0.929	0.617	0.617	0.763	0.740	0.927	0.894	0.586	0.936	0.769
FIM									0.815	0.464	0.513	0.639	0.708	0.847	0.856	0.580	0.814	0.619
FRF										0.674	0.681	0.781	0.616	0.938	0.807	0.616	0.923	0.685
GPB											0.633	0.845	0.626	0.552	0.491	0.337	0.644	0.658
ITL												0.568	0.679	0.570	0.636	0.782	0.794	0.574
JPY													0.629	0.730	0.683	0.401	0.740	0.773
MYR														0.564	0.815	0.607	0.750	0.778
NLG															0.829	0.616	0.911	0.676
SEK																0.672	0.884	0.782
SGD																	0.740	0.490
XEU																		0.780
ZAR																		

As it can be seen, currencies in the Euro zone were highly correlated during 1996. This is due to the fact that those currencies were part of the European Monetary System so their relative oscillations were controlled by their respective Central Banks. Surprisingly, a similar strong correlation has been found between Canadian Dollar, Japanese Yen and British Pound. This group even though it was not controlled exhibited the same type of behaviour. The results form Table 7 allows us the to define a measure of the degree of correlation between exchange currency rates time series.

Furthermore, it is possible by comparing the *%recur* to show the existence of correlation in time between several time series. For example, Fig. 71 shows that the Belgium Franc was correlated during all the year to the Euro; for the case of the Italian Lira, it can be seen, Fig. 72, a different correlation with the Euro before (points) and after (crosses) its entrance in the EMS; whereas for the Swiss Franc it is possible to establish two periods (January-March, left points) and (May-December, rigth points) with a transition between them (crosses), see Fig. 73. Hence, we can establish correlation periods between several exchange rates time series. This will help

us in two ways, first, it is possible to combine both series, when correlated, to improve our degree of precision in forecasting.

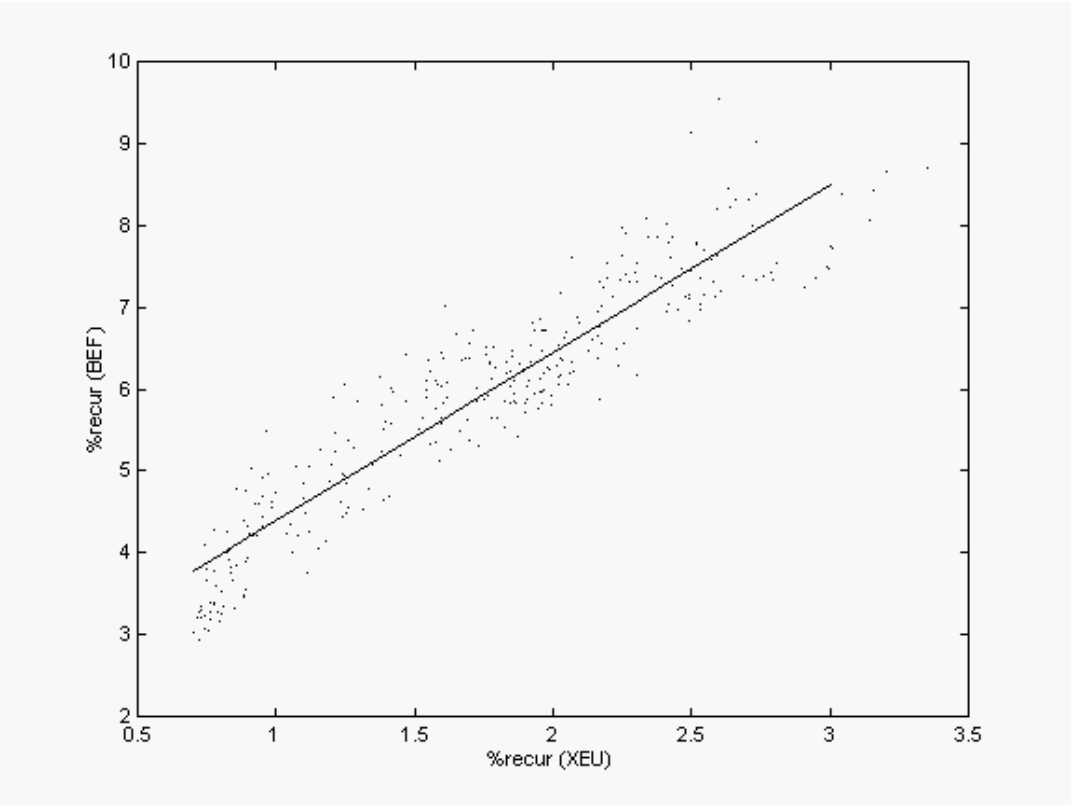

Figure 71. Plot of the *%recur* for the Euro and the Belgium Franc.

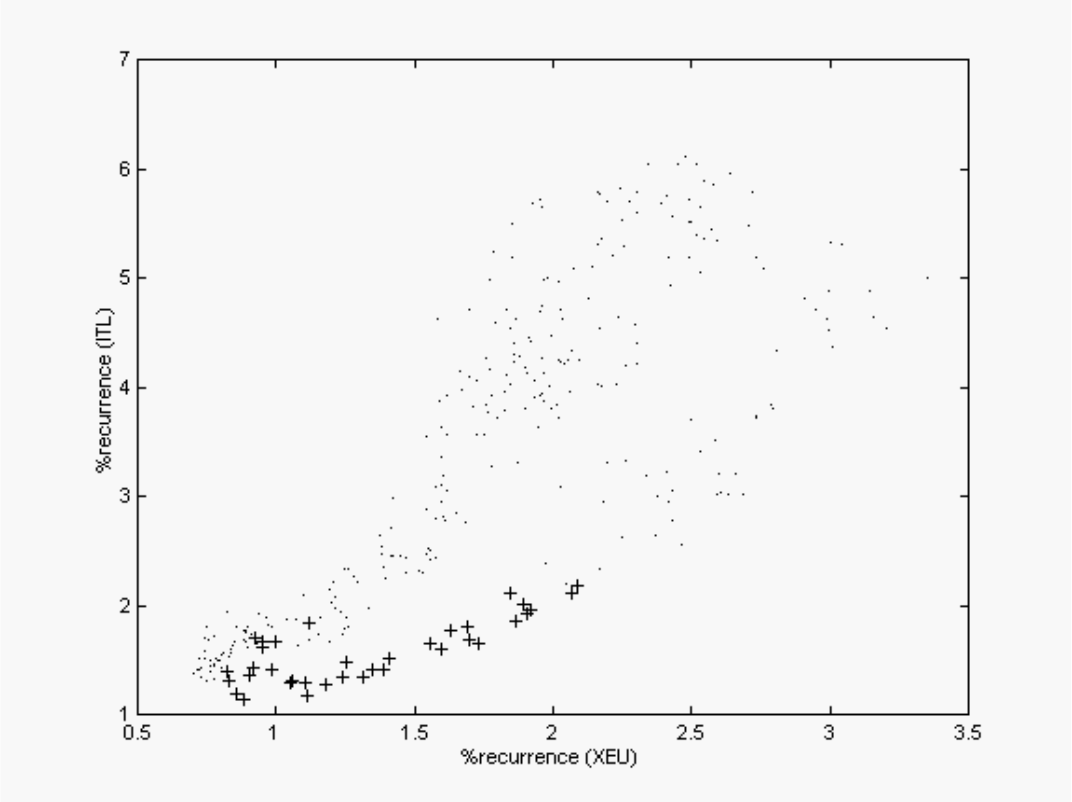

Figure 72. Plot of the *%recur* for the Euro and the Italian Lira. Blue values before the entrance in the EMS, green values after the entrance.

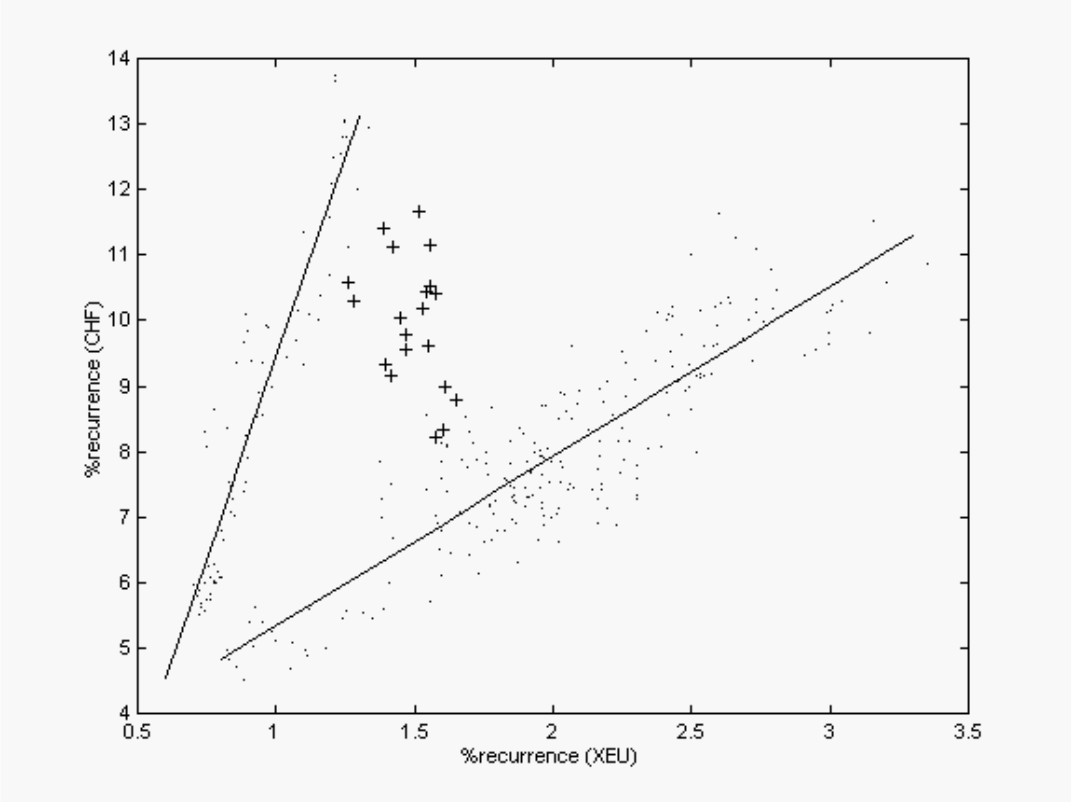

Figure 73. Plot of the *%recur* for the Euro and the Swiss Franc. Blue values for the period between January and March, green for the transition period and red for the last part of the year.

6. Conclusions

The Rescaled Range analysis (R/S analysis), through the calculation of the Hurst exponent, shows that there is persistence, $H>0.5$, in all time series, but this technique does not help in distinguishing between them.

The power spectral density shows a scaling behaviour typical of financial time series. The maximum exponent obtained is for the German Mark exchange rate time series and the minimum corresponds to the Singapore Dollar exchange rate time series. Again this methods does not help to classify the exchange rate clearly.

The non saturation in the space time separation plot shows that the time series may be nonstationary or that it has not been sampled long enough, and hence, the application of the surrogate data test that assumes a stationary Gaussian linear process as a null hypothesis is not adequate.

Rappresenting the data using the Recurrence Plot (RP) allows to individuate four different structures: constant RP over time, high recurrence during the first 6 months or in the last 6 months, aeroplane structure. Evidently that some time series are on the borderline of these classification schemata and to classify them in one group or in another is rather subjective. Individual analysis of high frequency currency exchange rates from 1996 using Recurrence Quantification Analysis (RQA) has allowed us to establish a correlation between changes in the system and external factors. The fact that there are considerable variations in the measure of

%determinism during the year may be indicative of changes in our forecasting potential. Our current research is oriented along these lines to quantify further this effect.

In order to compare the results obtained for all the time series we have calculated the RP on a weekely basis shifted by a day, and we have generated 305 values for each of the five RQA variables (*%det, %recurr, entropy, trend, 1/max*$_{line}$) as a function of time, using the same state space reconstruction parameters. We have plotted *%recur* of each time series as a function of *%recur* of the other currencies and used the regression coefficient to define a new measure of nonlinear correlation. We observe that in some cases, for example *%recur*(BEF) with respect to *%recur*(XEU) there is a strong correlation (linear behaviour) whereas in other cases, for example *%recur*(ITL) with respect to *%recur*(XEU), we have something like a phase transition from one line to onother in correspondence of the entrance of the Italian Lira in the Euro zone.

The results show, as expected, the high correlation between the currencies that were part of the Euro zone in 1996. This can be considered as a confirmation of the technique, but also an unexpected strong correlation between the Japanese Yen, the Canadian dollar and the British Pound. Furthermore our method has been found that, in the Euro zone, the most correlated currencies to the ECU, in 1996, were the Spanish Peseta and the Belgium Franc. However, the higher correlation coefficient corresponds to the Danish Krone, which even though it was not in the Euro zone had the typical behaviour of one of the currencies inside. At this point we have to remember that in 1996 Italian Lira and Finnish Markka entered in the Euro zone only at the end of the year i.e. respectively the 25 of November and 14 of October and data on Portuguese Escudo exchange rates were not available.

Finally, by inspection of the different plots, it is possible to observe changes in the degree of correlation during the year. One ingredient in a good investment strategy is the selection of groups of assets that tend not to move in the same direction together. Diversifying holdings helps investors to minimise the portfolio's risk against unexpected fluctuations in the market by investing in non correlated markets. This technique may find an application on this problem.

Acknowledgements.

The authors gratefully acknowledge the Director of the LIUC Library Mr. Piero Cavaleri who purchased the data sets and help us in the historical research and, also, to Olsen & Associates, which are the source of the data sets analysed. Furthermore, we would like to acknowledge Prof. Giorgio Perderzoli for his useful comments on a preliminary version of the manuscript.

References

Abarbanel, H.D.I., *Analysis of Observed Chaotic Data*, 1996, Springer-Verlag, New-York.

Andreadis, I., 2000, Self-criticality and stochasticity of an S&P 500 index time series. *Chaos, Solitons and Fractals* **11**, 1047-1059.

Badii, R., Broggi, G., Derighetti, B., Ravani, M., Ciliberto, S., Politi, A. and Rubio, M. A., 1988, Dimension increase in filtered signals. *Phys. Rev. Lett*. **60**, 979-982.

Bak, P. and Chen, K., 1991, Self-organized criticality. Scientific American 264, 26-33.

Breeden, J. L and N. H. Packard, 1994, A learning algorithm for optimal representation of experimental data, *Int. J. of Bifurcations and Chaos* **4**, 311- 326.

Brock, W. A., Hsieh, D. A., LeBaron, B., 1991, Nonlinear dynamics, chaos and instability:Statistical theory and Economic evidence. MIT Press, Massachusetts, MA.

Broomhead, D. S. and G. P. King, 1986, Extracting qualitative dynamics from experimental data, *Physica D* **20**, 217-236.

Burden, R. L. and Faires, J. D., *Numerical Analysis* , 3rd Ed., PWS, Boston.

Cannon, M. J. Percival, D. B., Caccia, D. C., Raymond, G. M. and Bassingthwaighte, J. B., 1997, Evaluating scaled windowed variance methods for estimating the Hurst coefficient of time series. *Physica A* **241**, 606-626.

Cao, L. ,1997, Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D* **110**, 43-50.

Cao, L., 2001, Method of False Nearest Neighbors, in *Modelling and Forecasting Financial Data*, A. S. Soofi and L. Cao (Eds.), Kluwer, Boston.

Casdagli, M., 1989, Nonlinear prediction of chaotic time series. *Physica D* **35**, 335-356.

Casdagli, M., Eubank, S., Farmer, J. D., Gibson, J., 1991, State space reconstruction in the presence of noise. *Physica D* **51**, 52-98.

Davies, M., 1994, Noise reduction schemes for chaotic time series, *Physica D* **79**, 174-192.

Diks, C., 1999, *Nonlinear Time Series Analysis: Methods and applications*.World Scientific, Singapore.

Eckmann, J. P., 1981, Roads to turbulence in dissipative dynamical systems, *Rev. Mod. Phys*. **53**, 643-654.

Eckmann, J. P., Kamphorst, S. O. and Ruelle, D., 1987, Recurrence plots of dynamical systems, *Europhys. Lett*. **4**, 973-977.

Fraser, A. and Swinney, H., 1986, Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* **33**, 1134-1140.

Friederich, R., Peinke, J. and Renner, Ch., 2000, How to quantify deterministic and random influences on the statistics of the foreign exchange market. Phys. Rev. Lett. 84, 5224-5227.

Gao, J. and Zheng, Z., 1994, Direct dynamical test for deterministic chaos and optimal embedding of a chaotic time series. *Phys.Rev. E* **49**, 3807-3814.

Gilmore, R., 1998, Topological analysis of chaotic dynamical systems. *Rev. Mod. Phys*. **70**, 1455-1529.

Grassberger, P., 1983, Generalized dimension of strange attractors. *Phys. Lett A* **97**, 227-230.

Haken, H., 1983, *Synergetics: An Introduction*, Springer-Verlag, Berlin.

Hegger, R., Kantz, H., Schreiber, T., 1999, Practical implementation of nonlinear time series methods: The TISEAN package. *CHAOS* **9**, 413-. The software package is publicly available at http://www.mpipks-dresden.mpg.de/~tisean .

Hsieh, D. A., Chaos and nonlinear dynamics: Application to financial markets, 1991, *The Journal of Finance* **46**, 1839-1887

Hurst, H. E., 1951, Long-term storage capacity of reservoirs, *Trans. Am. Soc. Civ. Eng*. **116**, 770-779.

Jensen, H. J., *Self-Organized criticality*, 1998, Cambridge University Press.

Judd, K. and Mees, A., 1998, Embedding as a modeling problem. *Physica D* **120**, 273-286.

Kantz H., Shreiber T., *Nonlinear Time Series Analysis*, 1997, Cambridge University Press

Kennel, M. B., R. Brown and H. D. I. Abarbanel, 1992, Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A* **45**, 3403-3411.

Kondratieff, N. D., 1935, The long wave in economic life, *Review of Economic Statistics* **17**, 105-115.

Kostelich, E. J. and Schreiber,T., 1993, Noise reduction in chaotic time-series data: A survey of common methods. *Phys. Rev. E* **48**, 1752-1763.

Kuznets, S., 1973, Models of economic growth: Findings and reflections, *American Economic Review* **63**, 247-258.

Lorenz, H. W., 1993, *Nonlinear dynamical economics and chaotic motion*. Springer, New York.

Malkiel, B., 1990, A random walk down Wall Street, Norton, New York.

Mandelbrot, B. B., 1998, Fractals and Scaling in Finance: Discontinuity, Concentration, Risk. Springer, New York.

Mandelbrot, B. B., *The Fractal Geometry of Nature*, 1983, W. H. Freeman. New York.

Manetti, C., M.-A. Ceruso, A. Giuliani, C.L. Webber, Jr., and J.P. Zbilut ,1999, Recurrence quantification analysis as a tool for the characterization of molecular dynamics sumilations. *Physical Rev. E* **59**, 992-998

Mantegna, R. N. and Stanley, H.E., 1995, Scaling behaviour in the dynamics of an economic index. *Nature* **376**, 46-49.

Mantegna, R. N. and Stanley, H.E., 1996, Turbulence and financial markets. *Nature* **376**, 46-49.

Mees, A. I., Rapp, P. E. and Jennings L. S., 1987, Singular value decomposition and embedding dimension. *Phys. Rev. A* **36**, 340.

Mitchel, W. C., *Bussines Cycles: The Problem and its Setting*. 1927, National Bureau of Economic Research, New York.

Nicolis, G. and Prigogine, I, *Self-organization in Complex Systems*, 1977, Wiley, New York.

Orsucci, F., K. Walter, A. Giuliani, C.L. Webber, Jr., and J.P. Zbilut ,1999, Orthographic structuring of human speech and texts: linguistic application of recurrence quantification analysis. *Int. J. Chaos Theory Appl*. **4**, 21-28.

Osborne, M. F.M., 1959, Brownian motion in the stock market. *Oper. Res*. **7**, 145-173.

Packard, N., Crutchfield, J., Farmer, D. and Shaw, R., 1980, Geometry from a time series. *Phys. Rev. Lett*. **45**, 712-715.

Papaioannou, G. and Karytinos, A., 1995, Nonlinear time series analysis of the stock exchange: The case of an emerging market. *Int. J. of Bifurcations and Chaos* **5**, 1557-1584.

Peters, E. E., *Chaos and Order in the Capital Markets: a New View of Cycles, Prices and Volatility*, 1996, 2nd Edition, Wiley, New York.

Provenzale, A., Smith, L. A., Vio, R. and Murante, G., 1992, Distinguishing between low-dimensional dynamics and randomness in measured time series. *Physica D* **58**, 31-49.

Ruelle, D., 1990, Deterministic chaos: the science and the fiction. *Proc. R. Soc. Lond. A* **427**, 241-248.

Sauer, T., Yorke, Y and Casdagli, M., 1991, Embedology, *J. Stat. Phys.* **65**, 579-616.

Scheinkman, J. and LeBaron, B., Nonlinear dynamics and stock returns. *J. Business* **62**, 311-318.

Schreiber, T. and Schmitz, A., 2000, Surrogate time series, *Physica D* **142**, 346-382.

Schreiber, T. and Schmitz, A.,1996, Improved surrogate data for nonlinearity tests. *Phys. Rev. Lett.* **77**, 35-38.

Schreiber, T., 1998, Interdisciplinary application of nonlinear time series methods. Physics Reports.

Schuster H.G., *Deterministic Chaos: an introduction*, 1995, VCH

Shlesinger, M. F., Zaslavsky, G. M., and Klafter, J., 1993, Strange kinetics. *Nature* **363**, 31-37.

Stark,J., Broomhead, D.S., Davies, M.E. and Huke, J.1997, Takens embedding theorems for forced and stochastic systems. *Nonlinear Analysis* **30**, 5303-5314.

Strozzi, F. and Zaldivar, J. M., 2001, Embedding theory: Introduction and applications to time series analysis. In: Soofi, A. and Cao, L. (Eds.), Nonlinear deterministic modelling and forecasting of economic and financial time series, Kluwer Academic (in press).

Takens, F., 1981, in Dynamical Systems and Turbulence, Warwick 1980, vol. 898 of Lecture Notes in Mathematics, edited by A. Rand and L.S Young, Springer, Berlin, pp. 366-381.

Takens, F., 1996, The effect of small noise on systems with chaotic dynamics. In *Stochastic and Spatial Structures of Dynamical Systems*, S. J. van Strien and S. M. Verduyn Lunel, Verhandelingen KNAW, Afd. Natuurkunde, vol. 45, pp. 3-15. North-Holland, Amsterdam.

Theiler, J., 1991, Some comments on the correlation dimension of $1/f^a$ noise. *Phys. Lett A* 155, 480-493.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. D., 1992, Testing for nonlinearity in time series: the method of surrogate data. *Physica D* **58**, 77-.

Tong, H., *Nonlinear Time Series: a Dynamical System Approach*. 1990, Oxford University Press. Oxford.

Trulla, L.L, A. Giuliani, J.P. Zbilut, and C.L. Webber, Jr. (1996). Recurrence quantification analysis of the logistic equation with transients. *Phys. Lett. A* **223**, 255-26.

Waldrop, M., *Complexity: The Emerging Science at the Edge of Chaos*, 1993, Touchstone Books.

Webber Jr. C. L. and Zbilut, J. P., 1994, Dynamical assessment of physiological systems and states using recurrence plot strategies. *J. Appl. Physiol*. **76**, 965-973.

Whitney, H., 1936, Differentiable manifolds. *Ann. Math.* **37**, 645-680.

Wolf A., J. B. Swift, H. R. Swinne, J. A. Vastan, 1985, Determining Lyapunov exponents from a time series, *Physica D* **16**, 285-317.

Zak, M., Zbilut, J. P and Meyers, R. E.,1997, *From Instability to Intelligence:Complexity and predictability in Nonlinear Dynamics*, Springer-Verlag, Heidelberg.

Zbilut, J. P. and Webber Jr. C. L., 1992, Embeddings and delays as derived from quantification of recurrence plots. *Phys. Lett. A* **171**, 199-203