BANK OF GREECE

COMPUTATIONAL INTELLIGENCE IN EXCHANGE-RATE FORECASTING

Andreas S. Andreou
George A. Zombanakis

Working Paper

No. 49 November 2006

BANK OF GREECE
Economic Research Department – Special Studies Division
21, E. Venizelos Avenue
GR-102 50 Athens
Tel: +30210-320 3610
Fax: +30210-320 2432

www.bankofgreece.gr

Printed in Athens, Greece
at the Bank of Greece Printing Works.

ISSN 1109-6691

COMPUTATIONAL INTELLIGENCE IN EXCHANGE-RATE FORECASTING

Andreas S. Andreou
University of Cyprus
and
George A. Zombanakis
Bank of Greece

ABSTRACT

This paper applies computational intelligence methods to exchange rate forecasting. In particular, it employs neural network methodology in order to predict developments of the Euro exchange rate versus the U.S. Dollar and the Japanese Yen. Following a study of our series using traditional as well as specialized, non-parametric methods together with Monte Carlo simulations we employ selected Neural Networks (NNs) trained to forecast rate fluctuations. Despite the fact that the data series have been shown by the Rescaled Range Statistic (R/S) analysis to exhibit random behaviour, their internal dynamics have been successfully captured by certain NN topologies, thus yielding accurate predictions of the two exchange-rate series.

Keywords: Exchange - rate forecasting, Neural networks
JEL Classification: C530

Acknowledgements: We are indebted to Heather Gibson and Tassos Anastasatos for their helpful comments. The views expressed in the paper are those of the authors and not necessarily reflect those of the Bank of Greece.

Correspondence:
George. A. Zombanakis,
Economic Research Department,
Bank of Greece, 21 E. Venizelos Ave.,
102 50 Athens, Greece,
Tel.: +30 210 323 5809
Fax: +30 210 323 3025
Email: gzombanakis@bankofgreece.gr

1. Introduction

Removing the risk of uncertainty and implementing the most suitable policy measure to face challenging economic policy problems at both a micro and a macro level have rendered exchange-rate prediction an issue very much in demand. Such policies, however, entailing the use of frequent market interventions by the various central banks in their effort to affect exchange-rate developments, contribute to increasing the level of noise traced in the historical series involved. It is thought that this issue adds to the problems faced during the forecasting exercise, although, as Taylor (1995) points out, empirical evidence on the link between official intervention and exchange rate expectations is rather unclear.

The fact remains, however, that the relevant literature is full of cases in which models aiming at forecasting exchange-rates yield results which are either poor (Marsh and Power, 1996; Pollock and Wilkie, 1996; West and Cho, 1995) or difficult to interpret (Kim and Mo, 1995; Lewis, 1989). Some authors even conclude that there is no such thing as "the best forecasting technique" and that the method chosen must depend on the time horizon selected or the objectives of the policy-maker (Verrier, 1989). There are, in addition, cases in which high persistence but no long range cycles have been reported (Peters, 1994), something that leads to the conclusion that currencies are true Hurst processes with infinite memory. On the other hand, long-term dependence, which supports the idea that cycles do exist in exchange rates, has been found by other researchers (Booth et. al., 1982, Cheung, 1993; Karytinos et. al., 2000). Despite various attempts to settle such controversies (Pesaran and Potter, 1992), it is still not clear whether the source of the dispute lies with the differences with respect to sample size, noise level, pre-filtering processes etc. of the various data sets employed, or the variety of tests that have been used, or, even, a combination of these factors.

But the major problem associated with exchange-rate predictability was highlighted by Meese and Rogoff (1983) and refers to the failure of the structural models to outforecast the random walk model, due, among other things, to difficulties in modeling expectations of the explanatory variables. The exchange-rate predictability issue has attracted increasing attention in the literature for more than a decade now involving authors like De Grauwe et al. (1993), Frankel (1993) and recently, Clark and McCraken (2001), Kilian and Taylor (2003), and Engel and West (2004). Although most sources agree on the empirical failure of models to forecast exchange-rate movements, none of them is able to provide a satisfactory explanation concerning the reasons why this is so. Most, however, seem to agree on the possibility that expectations are much more complicated than what modern exchange-rate theories have specified (see e.g. Pilbeam, 1995), primarily because the rapid flow of information, as well as the shift in the demand and supply pattern significantly influence market movements (Mehta, 1995). Other sources like Faust et al. (2003) focus on the nature of the data used, while only very recently, Evans and Lyons (2005) have underlined the importance of non-public information for the reliability of exchange-rate predictions, showing that the forecasting performance of a micro-based model is superior to that of a standard macro model and a random walk. The requirement of an accurate exchange-rate prediction appears even more demanding in cases of emerging markets in which "the trend towards the adoption of a more flexible exchange rate regime is also found to be important" (Amato et al, 2005) in an environment in which "the modern period of floating rates has seen such wide swings of real exchange rates………" (Fisher, 2005)[1].

In face of the limited success of the empirical literature to interpret exchange-rate movements, researchers have resorted, quoting Taylor (1995), to the use of "recently developed sophisticated time-series techniques". Indeed, one such technique is that which relies on tracing chaotic behaviour in exchange rates, as well as that of artificial neural networks. These, being data-driven approaches, have been considered preferable to traditional, model-driven approaches used for forecasting purposes, on the grounds of our earlier evaluation in the context of the literature cited. The advantages of the method used in these papers, namely that of neural networks, are extensively analysed in sources like Kuo and Reitsch (1995), Kosko (1992), Patterson (1996) and Haykin (1994),

[1] Concerning the Greek drachma, in particular, a small number of studies have attempted to forecast its exchange rates versus major currencies: Karfakis (1991) has concentrated on the drachma/USD exchange rate, while Koutmos and Theodosiou (1994) and Diamandides and Kouretas (1996), have explored the predictability of the drachma with respect to a number of European currencies, the USD and the Japanese Yen. In all these cases, however, the analysis is conducted in the context of exchange-rate models, with all the drawbacks that such a choice might entail.

underlining the fact that Neural Networks do not depend on the philosophy and structure of econometric models, an issue which often becomes the cause of scientific controversies. Indeed, given that NNs are non-linear they can capture complex interactions among the input variables in a system, thus being very useful in cases in which either standard theory cannot conclude as to a specific model structure or occasions in which a mathematical formulation is very hard or impossible or even cases in which immediate response to environment changes is required[2]. This means that in comparison to multiple regression analysis, in particular, the advantages of NNs are overwhelming: There is no need to choose any model specification since NNs are designed to perform automatically the so-called estimation of input significance as a result of which the most significant independent variables in the dataset are assigned high synapse (connection) weight values while showing negligible weight values for irrelevant variables. Relieved, thus, of the constraints imposed by model structures, NNs, being adaptive, can be trained, without depending upon prior knowledge of any rules, to learn underlying relationships on the basis of a training data set even when such relationships are difficult to find and describe. Once trained to recognize such relationships NNs can generalise by processing information that only broadly resembles the original training data set, a very useful property indeed given that real world data are often noisy. Similarly, they can handle imperfect or incomplete data by providing a measure of fault tolerance while they can account for any functional dependency as the network has the ability to discover (learn or model) the nature of such a dependency by itself. Finally, thanks to their parallel architecture, NNs can achieve high computational rates while posing no conditions on the predicted variables; it can be a two-state output (e.g. True/False, Yes/No), a numerical value, one or more classes among *n*, etc.

Thus NNs, requiring much less human work than traditional statistical analysis thus contributing to development – time reduction and easiness to handle, caused an expansion of their application in the exchange-rate literature. Such studies (e.g. Mehta,

[2] Thus, neural networks have contributed a great deal in areas like Signal Analysis and Processing as well as Process Control and Robotics. In addition they have proved to be very useful in cases of Data Classification and Smoothing, Pattern Recognition, Image and Speech Analysis and Medical Diagnostics. Finally, their contribution to defence issues, as well as to stock market or exchange rate forecasting analysis for loan or credit solicitations and marketing orientation has been widely acknowledged.

1995; Steurer, 1995; Refenes and Zaidi, 1995; Karytinos et. al., 2000 and Andreou et al., 2002), lead to better results concerning exchange-rate forecasting compared to "conventional methods".

The present analysis moves along the lines of this last paper, namely that by Andreou et al. (2002), that succeeded in forecasting the developments of the Greek Drachma versus the U.S. Dollar, the Deutsche Mark, the French Franc and the British Pound with nearly 99% accuracy by using NNs trained with Kalman filtering and evolved by a dedicated genetic algorithm in terms of topology and size. In the case of the present paper we shall attempt a forecast of the U.S. Dollar and the Japanese Yen rates versus the Euro using essentially the same technique applied to a series of five-minutes observations covering the last three months of 2004 (approximately 15,000 observations). The extent to which our series exhibit cyclical behaviour is investigated via various forms of R/S analysis presented in the second section of this paper. The different topologies Multi-layer Perceptrons trained with the Back Propagation algorithm are described in the third section of this paper and the corresponding results are analyzed and discussed in section four. The main conclusion that may be drawn here is that the NN employed managed to learn the underlying dynamics of the exchange-rate developments and yielded successful results of above 98% accuracy. Finally, the conclusions drawn are presented in the fifth section of this paper.

2. Long-term dependence detection techniques: R/S analysis

2.1 R/S analysis background

The origin of R/S analysis relies on the "T to the one-half rule", that is, on the formula that describes the Brownian motion:

$$R = T^{0.5} \tag{1}$$

where R is the distance covered by a random particle suspended in a fluid and T a time index.

It is obvious that (1) shows how R is scaling with time T in the case of a random system, and this scaling is given by the slope of the log(R) vs. log(T) plot, which is equal to 0.5. When a system or a time series is not independent the following formula can be used instead of (1):

$$(R/S)_n = c\, n^H \tag{2}$$

where, $(R/S)_n$ is the Rescaled Range statistic measured over a time index n, c is a constant and H the Hurst Exponent, which shows how the R/S statistic is scaling with time.

The R/S method estimates the Hurst exponent by transforming (2) to:

$$\log (R/S)_n = \log(c) + H \log(n) \tag{3}$$

and H can be estimated as the slope of the log/log plot of $(R/S)_n$ vs. n.

The analytical procedure to estimate the $(R/S)_n$ values, as well as, the Hurst exponent by applying (3), is described in Appendix I.

The Hurst exponent takes values from 0 to 1 ($0 \le H \le 1$). Gaussian random walks or, more generally, independent processes, give $H = 0.5$. If $0.5 \le H \le 1$ positive dependence is indicated, and the series is called persistent or trend reinforcing. In terms of equation (1), the system covers more distance compared to a random one, in which case the series exhibits a long memory process with no characteristic time scale. This scale invariance together with the existence of a power law (the log/log plot) are the key characteristics of a fractal series. If, on the other hand, $0 \le H \le 0.5$, this indicates negative dependence yielding anti-persistent or mean-reverting behaviour[3]. In terms of equation (1), the system in this case covers less distance than a random series, which means that it reverses itself more frequently than a random process.

A Hurst exponent different from 0.5 may characterise a series as fractal. However a fractal series might be the output of different kinds of systems. A "pure" Hurst process is a fractional Brownian motion (Mandelbrot and Wallis, 1969), also known as "biased random walk" or "fractal noise" or "coloured noise", that is, a random series the bias of which can change abruptly but randomly in terms of direction or magnitude. When dealing, however, with historical data the problem of distinguishing between the above alternatives becomes more difficult due to the existence of noise. In fact, given that most

[3] Only if the system under study is assumed to have a stable mean.

series are contaminated by either additive or dynamical (system) noise, the problem boils down to distinguishing between fractal noise and noisy chaos in most cases, including those dealing with financial data.

Although in general, fractal noises have no discernible cycles, in practice and in a certain time scale, fractional Brownian motion may exhibit a finite memory effect, which is usually a statistical artifact due to the limited length of the series, examined. In this case, fractal noise can be distinguished from a chaotic alternative by examining whether the cycle is independent of the time scale used, in which case it indicates the (noisy) chaos alternative.

Detection of cycles and estimation of their length can be accomplished by the use of the V-statistic (Hurst, 1956; Peters, 1994) defined as:

$$V_n = (R/S)_n / n^{0.5} \qquad (4)$$

The V_n vs. log(n) plot, gives a flat line for an independent random process and an upwardly sloping curve in the case of persistent series. The existence of a cycle and its length can be discerned (even in the presence of noise) from the “break-point” in this plot occurring when V_n reaches a peak and then flattens out, an indication that the long-memory process has dissipated.

2.2 R/S analysis application problems

The main problem in applying R/S analysis is the technical evaluation of the H exponent and, specifically, its statistical significance in comparison to a random null. Peters (1994) shows that under the Gaussian null, a modification of a formula developed by Anis and Lloyd (1976) allows for hypothesis testing by computing $E(R/S)_n$ and E(H), the expected variance of which will depend only on the total sample size N, as Var (H) = 1/N. However, if the null is still iid (independent and identically distributed) randomness but not Gaussianity, the formal hypothesis testing is not possible. To overcome this problem we used bootstrapping (Efron, 1979) to assess the statistical significance of the H exponents of our series, against both the Gaussian and the iid random null hypotheses.

To test against the Gaussian random null, we calculate the H exponent using 5,000 random shuffles of a Gaussian random surrogate, having the same length, mean and variance with our return series. The output is then compared to the test statistic i.e. the actual H exponent of the series. If the latter is found to be greater than 0.5, and persistence of the series is possible, then the null hypothesis tested is formed as : $H_0 : H = H_G$ with the alternative being: $H_1: H > H_G$. The significance level of the test is constructed as the frequency with which the pseudostatistic (the H_G estimate) from the Gaussian shuffles is greater than or equal to the actual statistic for the tested (unshuffled) data (Noreen, 1989). The null hypothesis is rejected if the significance level is smaller than the conventional rejection levels of 1%, 2,5% or 5%.

When the actual H statistic is found to be lower than 0.5 and anti-persistence is possible the null can be formed again as: $H_0 : H = H_G$ but the alternative this time is $H_1 : H < H_G$. In this case, the significance level of the test is constructed as the frequency with which the pseudostatistic H_G is smaller than or equal to the actual statistic and the null is rejected if the significance level is smaller than the conventional rejection levels of 1%, 2.5% or 5%.

To test against the iid null, the same procedure is followed but this time we randomize the series tested to produce 5,000 iid random samples having the same length and distributional characteristics as the original series. In this case, rejection of the null means that the actual H exponent calculated using the original series is significantly different (greater or smaller depending on the hypothesis tested) from the one calculated using an iid random series.

A second problem is related to the sensitivity of R/S analysis to short-term dependence, which can lead to unreliable results (Anis and Lloyd, 1976; Aydogan and Booth, 1988; Haubrich and Lo, 1980; Lo, 1991; Lo and Mackinlay, 1988; Milonas et al. 1985). Peters (1994), shows that Autoregressive (AR), Moving Average (MA) and mixed ARMA processes exhibit Hurst effects, but once short-term memory is filtered out by an AR(1) specification, these effects cease to exist. By contrast, ARCH and GARCH models do not exhibit long-term memory and persistence effects at all. Hence, a series should be pre-filtered for sort-term linear dependence before applying the R/S analysis.

2.3 The modified R/S analysis

An alternative way to account for short-term dependence is to use the R/S test statistic modified by Lo (1991). In Lo's modification, short-range dependence is incorporated into the partial sum variance estimator in the denominator of the classical Mandelbrot´s (1972) R/S statistic as:

$$Q_q = \hat{s}_N^{-1}(q)\left[\max_{1\le k\le N}\sum_{t=1}^{k}(X_t - \overline{X}_N) - \min_{1\le k\le N}\sum_{t=1}^{k}(X_t - \overline{X}_N)\right] \tag{5}$$

where

$$\hat{s}_N^2(q) = \hat{s}_x^2 + 2\sum_{t=1}^{q} w_t(q)\hat{g}_t, \qquad w_t(q) = 1 - \frac{t}{q+1}\ , \ q<N \tag{6}$$

$$\hat{s}_N^2(q) = \frac{1}{N}\sum_{t=1}^{N}(X_t - \overline{X}_N)^2 + \frac{2}{N}\sum_{t=1}^{q} w_t(q)\left\{\sum_{l=k+1}^{N}(X_l - \overline{X}_N)(X_{l-k} - \overline{X}_N)\right\} \tag{7}$$

and $\hat{s}_x^2$ and $\hat{g}_t$ being the variance and autocovariance estimators of X.

According to (6) and (7), if $\{X_t\}$ is subject to short-range dependence, the estimator $\hat{s}_N(q)$ involves the sums of squared deviations of X and its weighted autocovariances up to lag q.

This test, unlike classic R/S analysis described above, does not rely on subsample analysis. The test's null is short-term dependence, which operationally is defined by Lo as a "strong-mixing" process, a concept introduced by Rosenblatt (1956)[4] aiming at deriving the asymptotic distribution of Q_q.. Lo shows that under certain conditions which place restrictions on the maximal moments, the degree of distributional heterogeneity and the maximal degree of dependence in $\{X_t\}$, the statistic $V_q = N^{-(1/2)} Q_q$ converges to the range of a "Brownian bridge" on the unit interval, a well-defined random variable with mean $(\pi/2)^{(1/2)}$, variance $\pi^2/6-\pi/2$ and a positively skewed distribution function. The critical values of the test derived by the asymptotic cumulative distribution function are given in Table 1.

[4] Strong mixing requires that the maximal dependence between two events becomes trivially small as their separation time increases without bound.

Table 1: Asymptotic critical values of the modified R/S statistic

Probability level	0.5%	2.5%	5%	10%	90%	95%	97.5%	99.5%
Critical value	0.721	0.809	0.861	0.927	1.620	1.747	1.862	2.098

In terms of a brief evaluation, the main advantage of the test is that it allows for formal statistical testing and is robust to serial correlation and some forms of non-stationarity. It is specifically designed to distinguish between weakly dependent processes (e.g. ARMA) and strongly dependent processes (e.g. fractionally integrated [ARFIMA] models [Granger and Joyeux, 1980; Hosking, 1981]). Notice that the main characteristic of these strongly dependent processes is the slowly decaying autocorrelation functions and non-periodic cyclical patterns. Moreover, the test's null is wide enough to include models with longer-term correlations like the stochastic models of persistence proposed by Campbell and Mankiw (1987), Fama and French (1988) and Poterba and Summers (1988). The test, however, unlike the classic R/S analysis, is not able to specify the cycle length of the series tested, while there are, in addition, certain shortcomings related to the test "per se". Lo (1991) shows that there are forms of short-term dependence violating the assumptions of the test's null[5]. He also reports low power of the test against chaotic processes like the "tent-map", a long-range dependent process with very low autocorrelation. Hiemstra and Jones (1995), in addition, find that right and left-tailed bootstrapped critical values of the modified R/S statistic fall below their asymptotic counterparts, resulting in higher right-tailed and lower left-tailed rejection rates. According to their analysis, this is due to the test sensitivity to moment condition failure, i.e. to the magnitude of the maximal moment of their series, which is less than 4. Brock and de Lima (1995), in addition, use Monte Carlo simulation to find that the sampling distribution of the test is shifted to the left, relatively to the asymptotic distribution. A final problem concerning the application of R/S analysis is related to the

[5] For example, the test has no power against processes with maximal moments less than 4 violating the moment condition of the test, or the first difference of a stationary process violating the heterogeneity condition.

sensitivity of the test to the truncation lag-parameter q in equation (6). Lo (1991) employs Monte Carlo simulations to assess the power of the test, which declines with increasing q and decreasing sample size. In fact, even for sample sizes of N=1000, the empirical rejection rates were much lower than nominal sizes for q values exceeding $N^{1/3}$.

Little is known about the optimal choice of q, although Andrews (1991) suggests a data dependent formula given by:

$$q_N = 1 + \mathrm{INT}\left[\left(\frac{3N}{2}\right)^{\frac{1}{3}} . \left(\frac{2\hat{r}}{1-\hat{r}}\right)^{\frac{2}{3}}\right] \tag{8}$$

where, N is the data length and $\hat{\mathbf{r}}$ the estimated first order autocorrelation coefficient of the data However, the truncation lag given by this formula is optimal only for an AR(1) data generating process.

Although R/S analysis combined with bootstrapping for assessing the statistical significance of the H exponent provides a powerful tool for detecting persistent behaviour and long-term cycles, we have decided to cross – check our findings by employing, in addition, the modified R/S statistic using different q-lengths equal to q_n=INT[$N^{/4}$], INT[$N^{1/3}$], INT[$N^{1/2}$], as well as the q values derived from the data-dependent formula in (10). To assess the results of the test we have used the asymptotic and bootstrapped critical values of our series, the latter based on the test statistics derived from 5000 time-scrambled shuffling of our data, thus producing iid series consistent with the test's null and robust to violations of its moment condition.

2.4 Data and methodology

The US Dollar / Euro and the Yen / Euro exchange-rates are expressed in terms of the first differences of natural logarithms, their series composed of five-minute observations, covering, approximately, the last three months of 2004 thus yielding a total of 15,000 observations. Following a statistical description of the two time series we resort to using R/S analysis to find the value of the Hurst exponent and check the series for cycles by means of V-Statistics analysis and bootstrapping techniques. The tests conclude with the use of the modified R/S analysis to evaluate the results obtained by the original R/S tests.

Table 2 presents a selection of statistics for the U.S. Dollar / Euro series. It seems that the series displays a very low degree of asymmetry, as opposed to a rather pronounced kurtosis value (Figure 1).

Table 2: U.S. Dollar / Euro Series Statistics

Sample Size	14999
Average	0.0000024
Median	0.0000095
Variance	0.00013
Standard Deviation	0.01155
Minimum	-0.99715
Maximum	1.00028
Range	2
Lower Quartile	-0.000068
Upper Quartile	0.0000747
Skewness	0.07312
Kurtosis	7496.8

Figure 2 plots the $\log(R/S)_n$ against time ($\log(n)$), following which the value of the Hurst exponent is calculated to be $0.452 < 0.5$, which points to a white noise explanation with no indications of a cyclical behaviour.

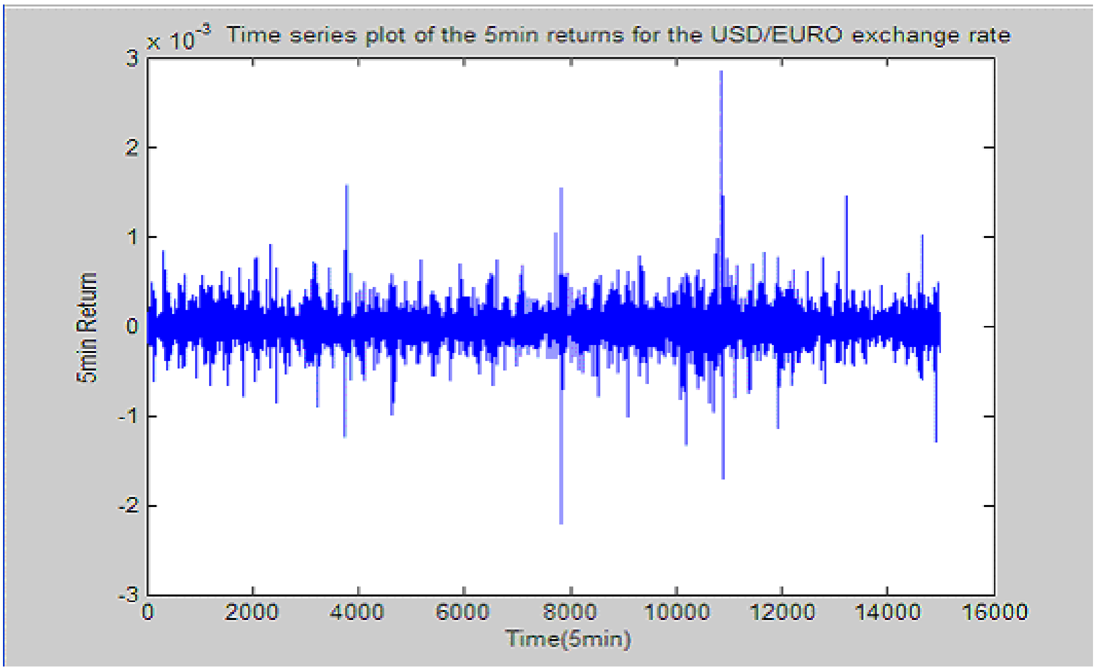

Figure 1: U.S. Dollar / Euro Logarithmic Time Series

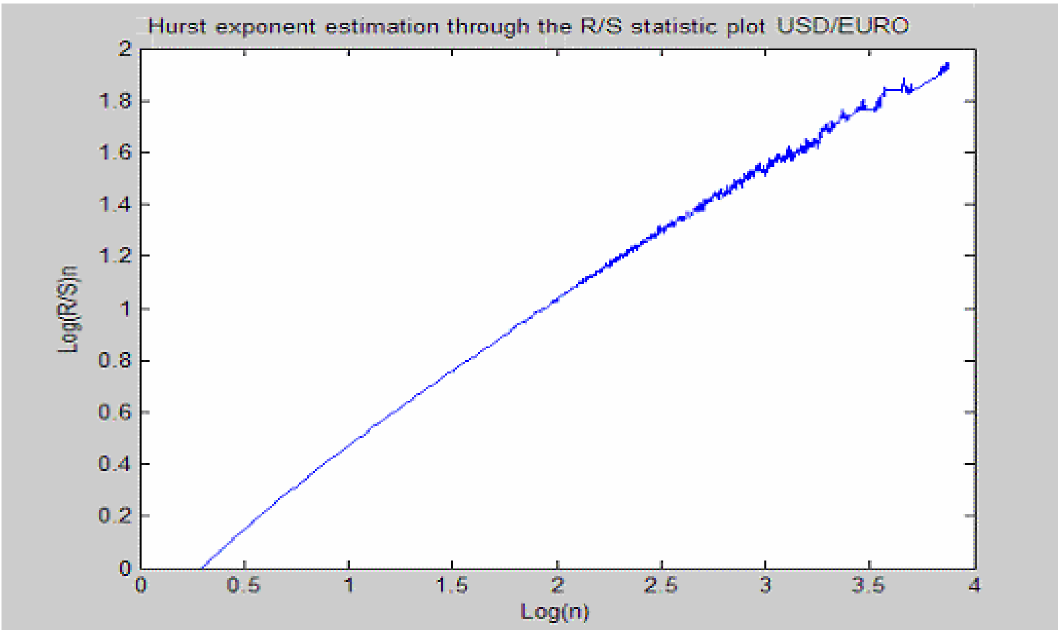

Figure 2: U.S. Dollar / Euro: Calculation of the Hurst exponent via the R/S plot

The results of the Gaussian and iid hypotheses tests reported in Table 3 support the R/S conclusions derived. In fact, the hypotheses against the Gaussian random and the random series with no long-term memory effect alternatives can not be rejected (significance levels of 48.6% and 23.76%, respectively).

Table 3: U.S. Dollar / Euro: Hurst exponent and Gaussian and iid Hypothesis Testing

A. R/S Analysis Results	$2<n<7500$
Hurst(H) Exponent	0.452
B. Bootstapping Results	$2<n<7500$
B1. Ho : $H > H_G$ (Gaussian Random Alternative)	
Mean H_G value	0.4731
Significance Level	48.6%
B2. Ho : $H > H_R$ (iid Random Alternative)	
Mean H_R value	0.4684
Significance Level	23.76%

Table 4 shows the V_q statistics for various q values, which turn out to support our results. In fact our bootstrapping experiment results point to accepting the hypothesis, given that the significance results obtained exceed 10% in all cases. Using the asymptotic values of Table 1 we conclude that the q values of the V_q estimates do not reject the test's short-range dependence below the significance level of 1%.

Table 4: U.S. Dollar / Euro: Modified R/S statistics and bootstrapped critical values

Q	Andrew's 1	N ¼ 11	N $^{1/3}$ 24	N ½ 122
Vq-Statistic	0.78	0.8016	0.8227	0.867
Bootstrapped Critical Values Significance Level				
1.0%	2	1.993	1.99	1.93
2.5%	1.867	1.865	1.855	1.83
5.0%	1.744	1.743	1.743	1.726
10.0%	1.612	1.616	1.613	1.597

All things considered, therefore, we can safely argue that the U.S. Dollar / Euro series is a random process with white noise characteristics.

Turning now to the Yen / Euro series, the basic statistics shown in Table 5 indicate just traces of asymmetry and kurtosis, while Figure 3 presents the time plot of this series.

Table 5: Yen / Euro Series Statistics

Sample Size	14999
Average	0.000000651
Median	0.00001581
Variance	0.0000000211
Standard Deviation	0.0001454
Minimum	-0.002018
Maximum	0.001545
Range	0.0035635
Lower Quartile	-0.0000795
Upper Quartile	0.00007964
Skewness	-0.35231
Kurtosis	10.270342

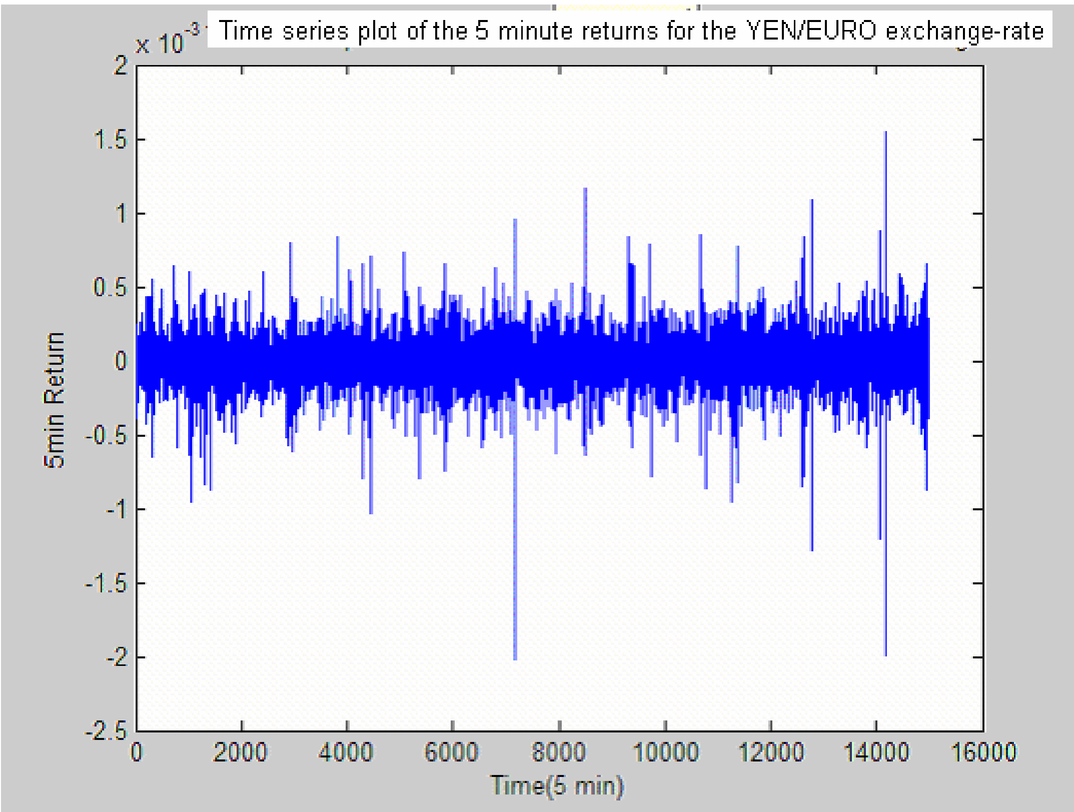

Figure 3: Yen / Euro Logarithmic Time Series

Figure 4 plots the $\log(R/S)_n$ against time, following which the value of the Hurst exponent is calculated to be 0.5121 > 0.5, which marginally allows for the possibility of a cyclical behaviour that can be traced using the V-statistic.

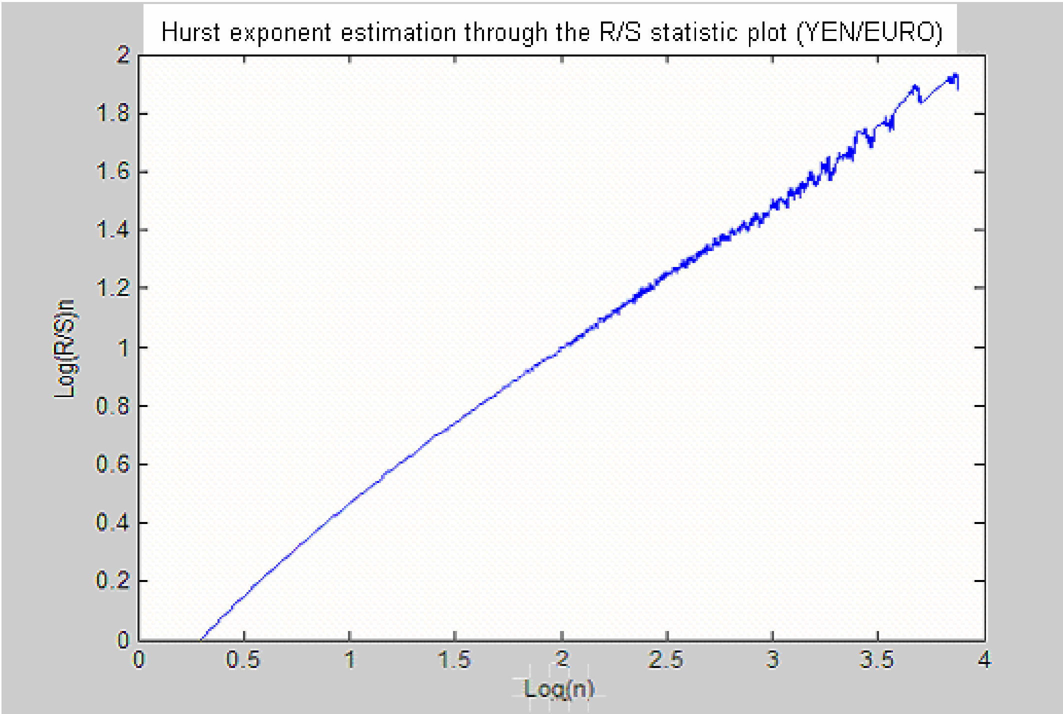

Figure 4: Yen / Euro: Calculation of the Hurst exponent via the R/S diagramme

Figure 5 depicts the V_q-statistic results, which do not point to the possibility of a cycle, unless we assume that a substantial difference between two consecutive V – values would indicate the presence of such a cycle. Such a difference has been traced when n = 2,992. Table 6, however, shows that the Hurst exponent value in this specific case is lower than 0.5, which does not reveal any sort of cyclical behaviour. These findings are supported by the Gaussian and iid hypotheses tests, which yield bootstrapped critical values with significance levels much higher than 5%. Besides, the Hurst exponent calculated does not reject the iid random alternative something that reinforces our conclusion.

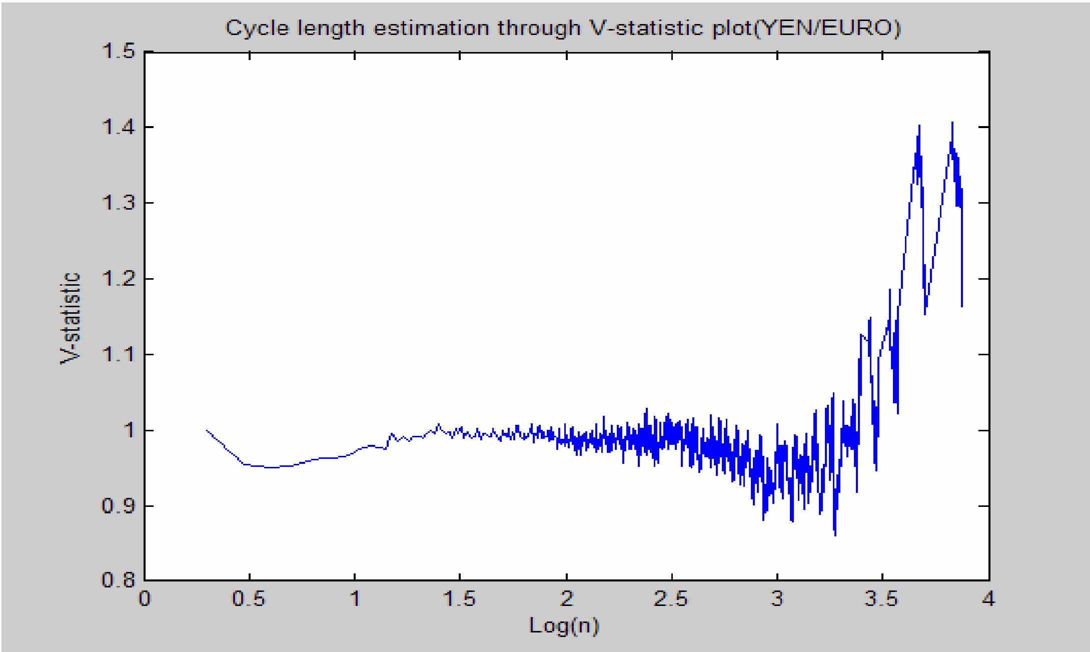

Figure 5: Yen / Euro: Testing for cyclical behaviour using V-Statistic diagramme.

Table 6: Yen / Euro: Hurst exponent and Gaussian and iid Hypothesis Testing

A. R/S Results	2<n<2992 (Cycle length)	2<n<7500
Hurst(H) exponent	0.496	0.5121
B. Bootstrapping Results	2<n<2992 (Cycle length)	2<n<7500
B1. Ho : $H > H_G$ (Gaussian random alternative)		
Mean H_G value	0.511	0.523
Significance Level	43.5%	23.5%
B2. Ho : $H > H_R$ (iid random alternative)		
Mean H_R value	0.519	0.5012
Significance Level	50.3%	63.3%

The modified R/S statistics seem to confirm our previous findings. Table 7 presents the V_q statistics for various q values the significance of which, as indicated by the bootstrapped critical values shows that the hypothesis couldn't be rejected even at a 10% significance level. The same applies when using the asymptotic critical values where the V_q values do not reject the test at a lower level than 10%.

Table 7: Yen / Euro: Modified R/S statistics and bootstrapped critical values

Q	Andrew's 1	N ¼ 11	N $^{1/3}$ 24	N ½ 122
Vq-Statistic	0.197	0.48	0.706	1.55
Bootstrapped critical values Significance level				
1.0%	2.01	1.98	1.99	1.96
2.5%	1.848	1.873	1.846	1.842
5.0%	1.744	1.751	1.739	1.72
10.0%	1.616	1.663	1.61	1.6

We can thus argue that despite initial indications pointing to the existence of cyclical behaviour in the Yen / Euro series, the results overall do not support such a conclusion.

3. Artificial Neural Networks

3.1 Basic NNs theoretical notions

This section is devoted to a brief presentation of the artificial neural networks methodology. This technique is a data driven approach, which is based on developing a "machine" composed of a number of interacting basic computational elements called neurons, distributed in layers and connected to each other. A network is trained through general-purpose algorithms based on time-series data and focusing on the computation of weight neuron connections in a feed-forward network to accomplish a desired input-output mapping. The learning phase of the computation procedure can be viewed as a high dimensional, non-linear, system identification problem. In a feed-forward Multi-Layer Perceptron (MLP) links from each neuron in the k^{th} layer are being directed to each neuron in the $(k+1)^{th}$ layer. Inputs from the environment enter the first layer and outputs from the network are manifested at the last layer. A typical m-d-1 MLP architecture is shown in Figure 6, which refers to a network with m inputs, d neurons in the hidden layer and one neuron in the output layer.

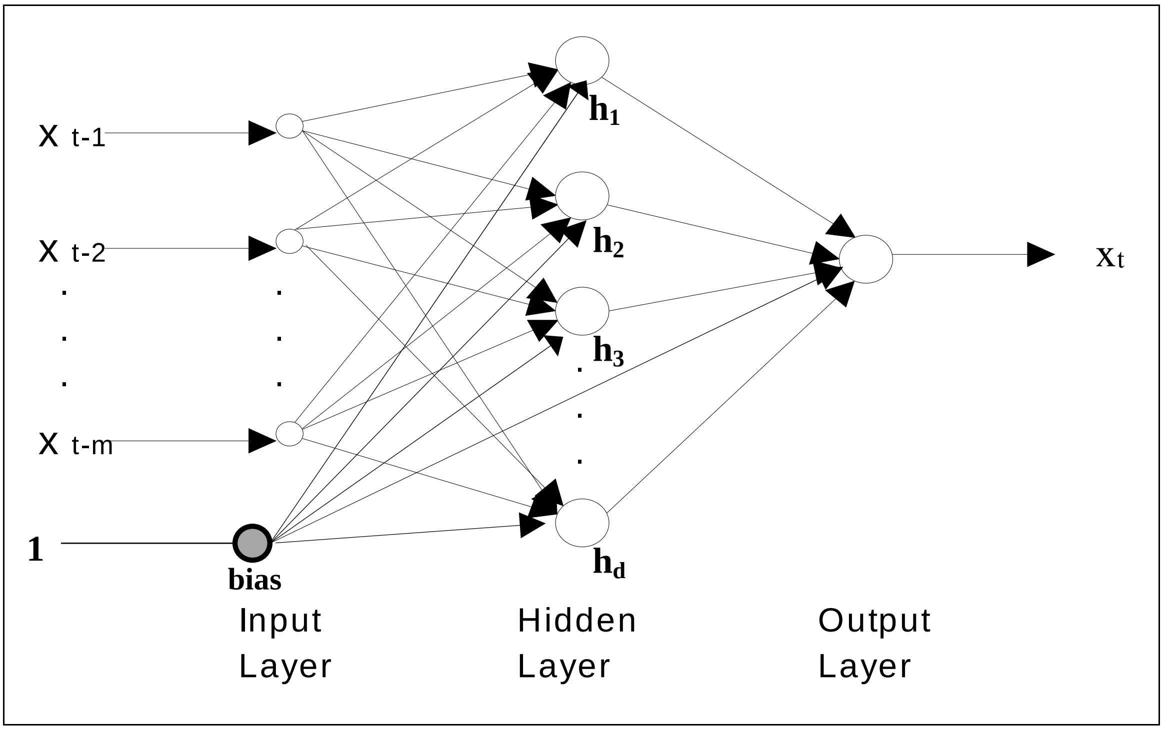

Figure 6: Typical MLP Neural Network Architecture with m input nodes, d nodes in the hidden layer and one output

Figure 6 shows, in addition, a special node at the end of the input layer called "bias". This node has a fixed input value of 1 and feeds into all the neurons in the hidden and the output layers, with adjustable weights as the other nodes. Its role is to represent the adjustable neuron threshold levels explicitly in the transfer function input. The nodal representation eliminates the need to treat the threshold as a special neuron feature and leads to a more efficient algorithm implementation (Azoff, 1994).

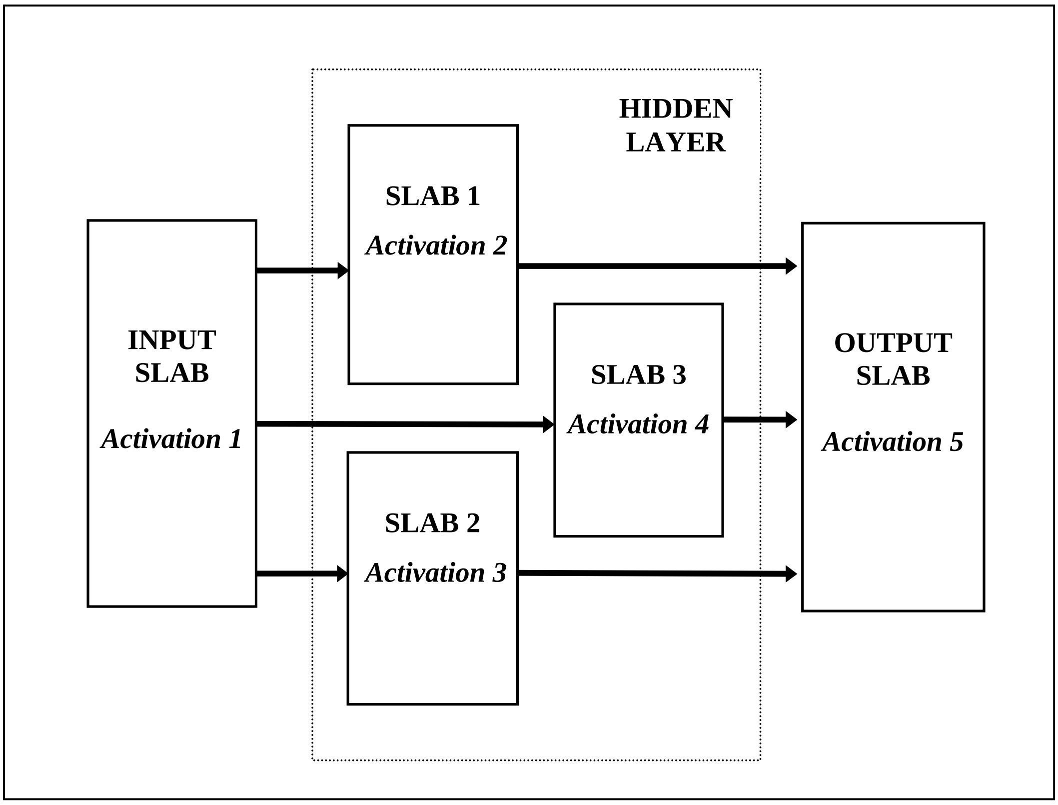

Figure 7: Multiply Activated MLP Neural Network Architecture

The networks used in the present paper were divided into three categories: The first one employs MLPs with a single hidden layer (category A), the second one includes MLPs with two successive hidden layers (category B) and the last one involves a Multiply Activated MLP (MAMLP – category C) which uses one hidden layer partitioned into three parallel sub-layers activated by a different function (Figure 7). Different topologies, as regards the number of nodes within the hidden layers, were implemented. In addition, variations of learning schemes were adopted, lying on different activation functions, such as:

$$\text{Logistic sigmoid}: \quad f(y) = (1 + \exp(-by))^{-1} \tag{9}$$

$$\text{Hyperbolic tangent}: \quad f(y) = (1 - \exp(-by)) * (1 + \exp(-by))^{-1} \tag{10}$$

$$\text{Gaussian}: \quad f(y) = \exp\left(-x^2\right) \tag{11}$$

Gaussian complement : $$f(y) = 1 - \exp\left(-x^2\right) \qquad (12)$$

where, $$y = \sum_{i=1}^{n} w_i x_i \qquad (13)$$

and x_i*'s* denote the input values of a node, while w_i's the real valued weights of edges incident on a node and *n* the number of inputs to the node from the previous layer. *b* is known as the steepness of equations (9) and (10). The input layer is linear, while the output uses the sigmoid function.

3.2 System design and implementation

The time series under study, x={x(t): 1 ≤ t≤ N} is divided into two sets: a training set x_{train}={x(t): 1 ≤ t ≤ T}, and a test set x_{test}={x(t): T < t ≤ N}, where N is the length of the data series. The training phase presents the x_{train} set to the network repeatedly until a certain level of convergence is achieved based on some error criterion. The generalization ability of the trained networks is then assessed by presenting the testing set, which is new data to the network not used during the training cycles, and producing forecasts. If the error criterion applied to the results obtained with the testing set is almost the same with that of the training set (possibly with slightly increased error figures) then generalization has been achieved. Otherwise, overfitting and memorization of the training data during the learning process is highly likely. The learning algorithm adjusts the weights in each repetition in order to minimize the diversion of the actual value from the predicted one. All networks developed have one output neuron, which yields the next sample (predicted value) in the time sequence. The training algorithm used is the well-known Error Back Propagation with a momentum term (Rumelhart and McLelland 1986; Azoff 1994).

The number of input neurons and the selection of the variables involved have been based on prior research on the topic that has led to the choice of the input set which exhibits the highest performance in terms of prediction accuracy. We used several alternative configuration schemes, as regards the number of hidden layers and the nodes within each layer, in order, both to achieve best performance and facilitate comparison between different network architectures. Every input variable is associated with one neuron in the input layer.

Determining the number of hidden layers and neurons in each layer can often be a very difficult task and possibly one of the major factors influencing the performance of the network. Too few neurons in a hidden layer may produce bias due to the constraint of the function space, which results in poor performance as the network embodies a very small portion of information presented. By contrast, too many neurons may cause over-fitting of the data, on one hand, and increase considerably the amount of computational time needed for the network to process data, on the other, something that will not necessarily lead to convergence. We therefore use a variety of numbers of neurons within one hidden layer, while in some cases a two-hidden-layer scheme is also developed in order to investigate whether performance may be improved.

The number of iterations (epochs) presenting the whole pattern set during the learning phase is also very important. We have allowed this number to vary during our simulations, since different network topologies, initial conditions and input sets, require different convergence and generalization times. The number of epochs our networks needed for convergence was 2000, while the learning and momentum coefficients (Rumelhart and McLelland 1986; Azoff 1994) were kept constant at the positive values of 0.3 and 0.1 respectively. One should be very cautious though when using a large number of epochs, as the network may over-fit the data thus failing to generalize. The problems of bias and data over-fitting can be overcome by evaluating the performance of each network using a testing set of unseen patterns (testing phase). This set does not participate during the learning process (see e.g. Azoff, 1994). If the network has actually learned the structure of the input series rather than memorizing it as previously mentioned, then it can perform well when the testing set is presented. Otherwise, if bias or over-fitting is really the case, performance will be extremely poor on these "new" data values. Architecture selection is generally based on success during the testing phase, provided that the learning ability was satisfactory.

Of the 15000 five-minute observations in our data set, 12,000 are included in the training set (80%) and 3,000 in the testing set (20%). The forecasting horizon is set to one step ahead.

3.3 Performance evaluation

Performance was evaluated using well known and widely used error measures, specifically the Normalized Root Mean Squared Error (NRMSE), the Correlation Coefficient (CC), the Mean Relative Error (MRE), the Mean Absolute Error (MAE) and the Mean Square Error (MSE). All these measures were evaluated on the testing set of data, that is, a set of pattern values that did not participate during the course of learning.

The CC measures the ability of the predicted samples to follow the upward or downward jumps in the original series. A CC value near 1 in absolute terms is interpreted as a perfect follow up of the original series by the forecasted one. A negative CC sign indicates that the forecasting series follows the same ups or downs of the original series with a negative mirroring, that is, with an 180^{o} rotation about the time-axis. When the original series moves up, the forecasting moves series down at the same time-period and vice versa.

The NRMSE indicates whether prediction is better than a simple mean forecaster. If NRMSE=0 then predictions are perfect; NRMSE=1 indicates that prediction is no better than taking x_{pred} equal to the x-mean.

The MRE shows the accuracy of predictions in percentage terms expressing it in a stricter way, since it focuses on the sample being predicted. It is also scale and unit independent. Thus, we are able to estimate prediction error as a fraction of the actual value, this making the MRE the more objective error measure among the three used.

The MSE, finally, is reported in order to have the error condition met by the Back Propagation algorithm, while the MAE shows the divergence between actual and predicted samples in absolute measures. The formulas for these prediction error measures are included in Appendix II.

4. Experimental results

The raw exchange-rate data are fed into various artificial neural networks selecting one every three, six or twelve observations to correspond to a forecasting horizon of 15, 30 and 60 minutes respectively. The results obtained for the U.S. Dollar / Euro exchange-rate series are presented in Table 8 with the most successful network results in bold characters. It should be noted that prediction success is assessed in the testing phase and that the priority given to the error metrics is the following: MRE first, NRMSE second and CC third.

It should be mentioned that cases (i) and (ii) use different activation functions for the hidden layers. More specifically, case (i) uses the Gaussian Complement (GauComp) for the first layer, the Hyperbolic Tangent (tanh) function for the second and the sine for the third. In case (ii), the first hidden layer uses the tanh, the second the Gaussian and the third the logistic sigmoid. All remaining topologies use the logistic sigmoid function for all layers.

Most of the results reported yield very satisfactory values in terms of all the criteria used. In fact, the lowest error values are observed in our final experiment, the architecture of which uses 12 input neurons and 2 hidden layers with 5 neurons each. Similar results but slightly inferior were achieved with the topology having 12 inputs and 5 neurons in a single hidden layer. Figure 8 compares the actual to the forecasted results during the training phase, using dark lines for the former and light for the latter, while Figure 9 depicts the testing phase results, which are very encouraging in terms of the errors and the correlation coefficient reported (very low MRE, CC close to 1 and NRMSE close to 0). We can safely argue, therefore, that the forecasting performance of a network architecture 12-5-5-1 is reliable enough to predict the U.S. Dollar / Euro rate.

Table 8: U.S. Dollar / Euro: Experiment Results

Network	Training Phase					Testing Phase				
	NRMSE	**MSE**	**CC**	**MRE**	**MAE**	**NRMSE**	**MSE**	**CC**	**MRE**	**MAE**
3-5-1	0.0786	0.045	0.9972	0.0014	0.1763	0.4754	0.042	0.9949	0.0013	0.1719
3-10-1	0.0651	0.0309	0.9979	0.0012	0.1508	1.3058	0.3165	0.9950	0.0041	0.5508
6-5-1	0.0735	0.0394	0.9974	0.0013	0.1626	0.6950	0.0892	0.9950	0.0019	0.2639
6-10-1	0.0839	0.0513	0.9977	0.0016	0.2073	0.5642	0.0652	0.9949	0.0016	0.2170
12-5-5-1	0.0187	0.0026	0.9998	0.00026	0.0343	0.1440	0.00035	0.9989	0.00037	0.0514
12-10-1	0.0787	0.0453	0.9974	0.0014	0.1885	0.5577	0.0525	0.9993	0.00014	0.1936
3-5-5-1	0.0288	0.0066	0.9996	0.0005	0.0663	0.5611	0.0194	0.9831	0.00096	0.1313
3-10-10-1	0.0323	0.00082	0.9995	0.00054	0.0708	1.0364	0.0661	0.9823	0.0018	0.2485
6-5-5-1	0.0235	0.0041	0.9998	0.00037	0.0486	0.5886	0.0576	0.9944	0.0016	0.2182
6-10-10-1	0.0607	0.0269	0.9982	0.00093	0.1202	1.1979	0.2418	0.9945	0.0034	0.4667
12-5-1	**0.0206**	**0.0031**	**0.9998**	**0.00023**	**0.0305**	**0.1367**	**0.0032**	**0.9987**	**0.00035**	**0.0481**
12-10-10-1	0.0269	0.0053	0.9996	0.00045	0.0584	0.1194	0.0024	0.9990	0.00031	0.0424
3-5-5-5-1	0.0358	0.0125	0.9992	0.00059	0.0778	1.3474	0.1118	0.9823	0.0024	0.3245
3-8-8-8-1	0.0283	0.0063	0.9996	0.00045	0.0589	1.1114	0.0760	0.9922	0.0019	0.2652
3-10-10-10-1	0.0321	0.0082	0.9995	0.00055	0.0715	1.0937	0.0736	0.9823	0.0019	0.2632
6-5-5-5-1	0.0285	0.0064	0.9996	0.00046	0.0599	0.9108	0.051	0.9814	0.0016	0.2161
6-8-8-8-1	0.0306	0.0069	0.9996	0.0005	0.0662	0.7245	0.0868	0.9940	0.0020	0.2748
6-5-8-10-1(i)	0.0240	0.00043	0.9997	0.00038	0.05	0.3403	0.018	0.9919	0.00087	0.1191
12-3-8-5-1(ii)	0.0204	0.0031	0.9998	0.00063	0.0396	0.2510	0.0074	0.9984	0.00054	0.0741

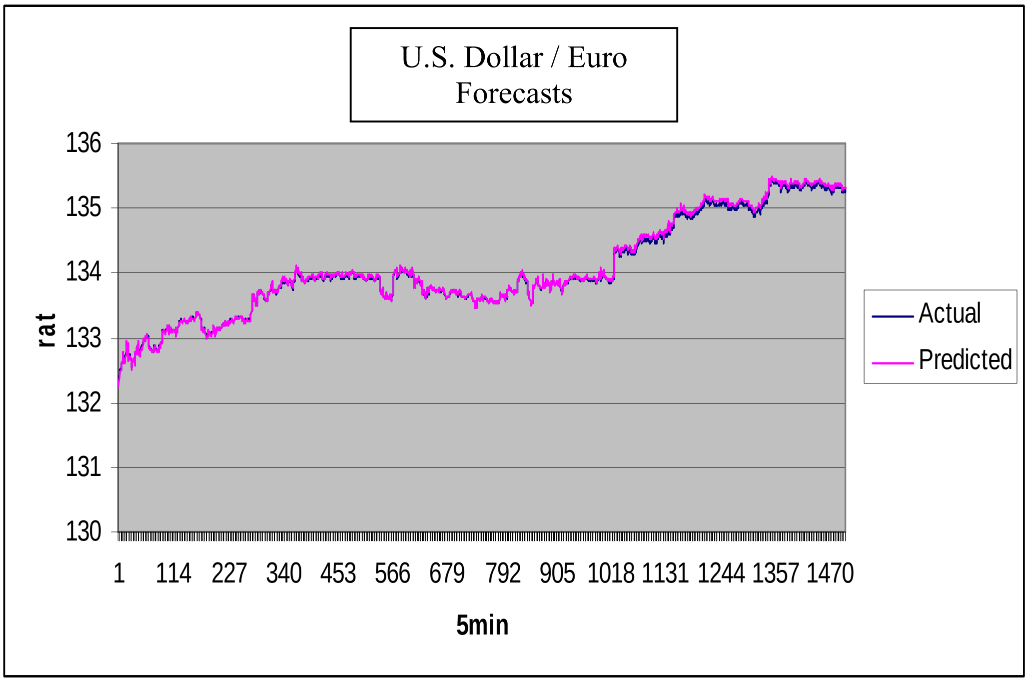

Figure 8: U.S. Dollar /Euro: Network Architecture 12-5-5-1 (Training Phase Data)

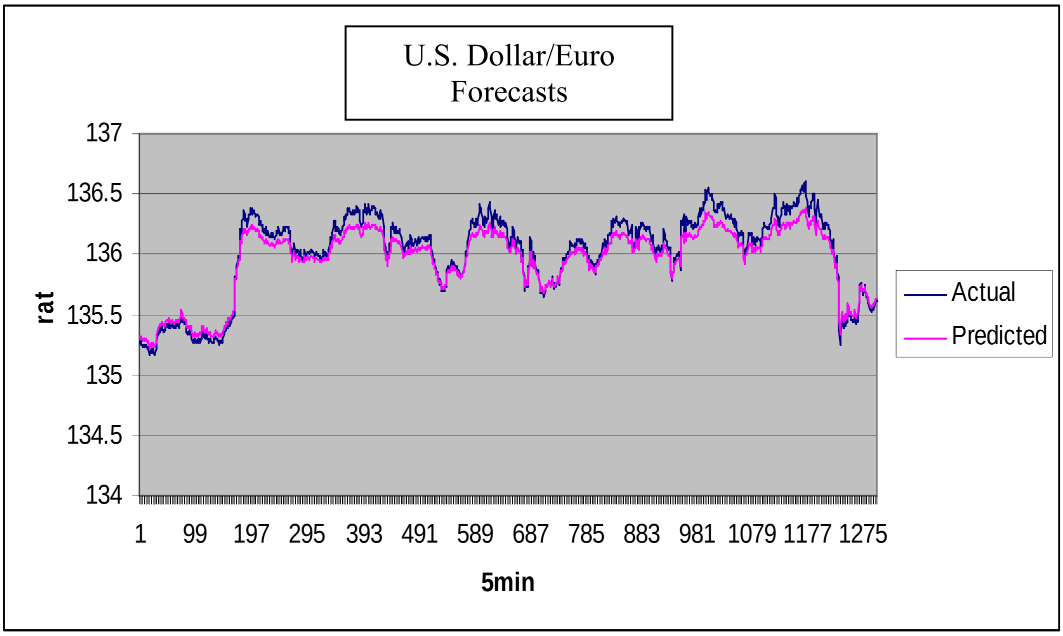

Figure 9: U.S. Dollar / Euro: Network Architecture 12-5-5-1 (Testing Phase Data)

The results for the Yen / Euro series are included in Table 9 as follows:

Table 9: Yen / Euro: Experiment Results

Network	Training Phase					Testing Phase				
	NRMSE	MSE	CC	MRE	MAE	NRMSE	MSE	CC	MRE	MAE
3-5-1	0.0934	0.0047	0.9958	0.00034	0.0472	0.1134	0.0051	0.9948	0.0004	0.0560
3-10-1	0.0970	0.0051	0.9454	0.00036	0.0495	0.1267	0.0063	0.9933	0.0004	0.0560
6-5-1	0.0701	0.0027	0.9978	0.00027	0.037	0.0966	0.0037	0.9957	0.00032	0.0441
6-10-1	0.0975	0.0051	0.9953	0.00037	0.05	0.1298	0.0066	0.9931	0.00041	0.0572
12-5-1	**0.0663**	**0.0025**	**0.9979**	**0.00036**	**0.0361**	**0.0778**	**0.0031**	**0.9972**	**0.00029**	**0.0401**
12-10-1	0.1010	0.0055	0.9951	0.00037	0.0504	0.1352	0.0072	0.9929	0.00042	0.058
3-5-5-1	0.1236	0.0083	0.9926	0.00044	0.0663	0.1370	0.0074	0.9919	0.00045	0.0627
3-10-10-1	0.1325	0.0095	0.9924	0.00045	0.0621	0.1210	0.0058	0.9930	0.00039	0.0538
6-5-5-1	0.1135	0.007	0.9973	0.00046	0.0626	0.1296	0.0066	0.9957	0.00048	0.0662
6-10-10-1	0.1103	0.0066	0.9960	0.00043	0.0585	0.1072	0.0045	0.9948	0.00036	0.05
12-5-5-1	0.0895	0.0043	0.9974	0.00034	0.0471	0.1036	0.0042	0.9960	0.00036	0.0495
12-10-10-1	0.1467	0.0116	0.9963	0.00058	0.0798	0.1616	0.0103	0.9954	0.00063	0.0869
3-5-5-5-1	0.1236	0.0083	0.9926	0.00045	0.0615	0.1293	0.0066	0.9929	0.00046	0.0635
3-8-8-8-1	0.1318	0.0094	0.9921	0.00046	0.0623	0.1871	0.0138	0.9889	0.0006	0.2826
3-10-10-10-1(i)	0.1053	0.006	0.9946	0.00038	0.0528	0.1235	0.006	0.9933	0.0004	0.0555
6-5-5-5-1	0.1566	0.0133	0.9910	0.00053	0.0728	0.1141	0.0051	0.9935	0.00038	0.0524
6-8-8-8-1(ii)	0.1328	0.0045	0.9925	0.00044	0.0605	0.1231	0.006	0.9930	0.00041	0.0575
6-5-8-10-1	0.1221	0.0081	0.9967	0.00049	0.0677	0.1208	0.0057	0.9951	0.00044	0.0605
6-15-15-15-1	0.1735	0.0163	0.9930	0.00067	0.0916	0.1540	0.0049	0.9933	0.00054	0.0749
12-3-8-5-1(ii)	0.1378	0.0103	0.9961	0.00053	0.0731	0.1319	0.0068	0.9954	0.00049	0.0674

Cases (i) , (ii) and (iii) use different activation functions for the hidden layers. In fact the first case uses a logistic function for the first hidden layer, a Hyberbolic Tangent for the second and a sine for the third. In the second case the first hidden layer uses a Gaussian function, the second a Gaussian Complement and the third a logistic sigmoid.

Finally, in the third case, the first hidden layer uses a Hyberbolic Tangent, the second a Gaussian and the third a logistic sigmoid.

In terms of a brief, overall assessment, our results for the Yen / Euro series seem to be slightly superior in terms of predictive ability compared to those derived for the U.S. Dollar / Euro series. The most successful experiment on the basis of the criteria used is the one involving the architecture 12-5-1, the results of which can be considered as rather encouraging (Figures 10 and 11).

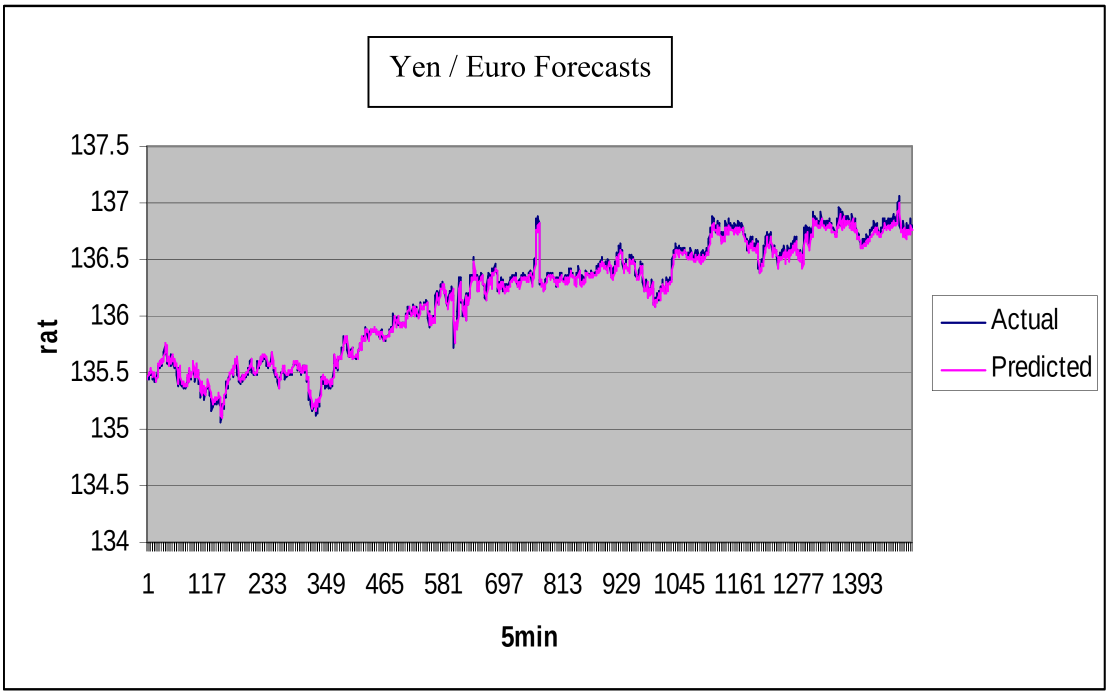

Figure 10: Yen / Euro: Network Architecture 12-5-1 (Training Phase Data).

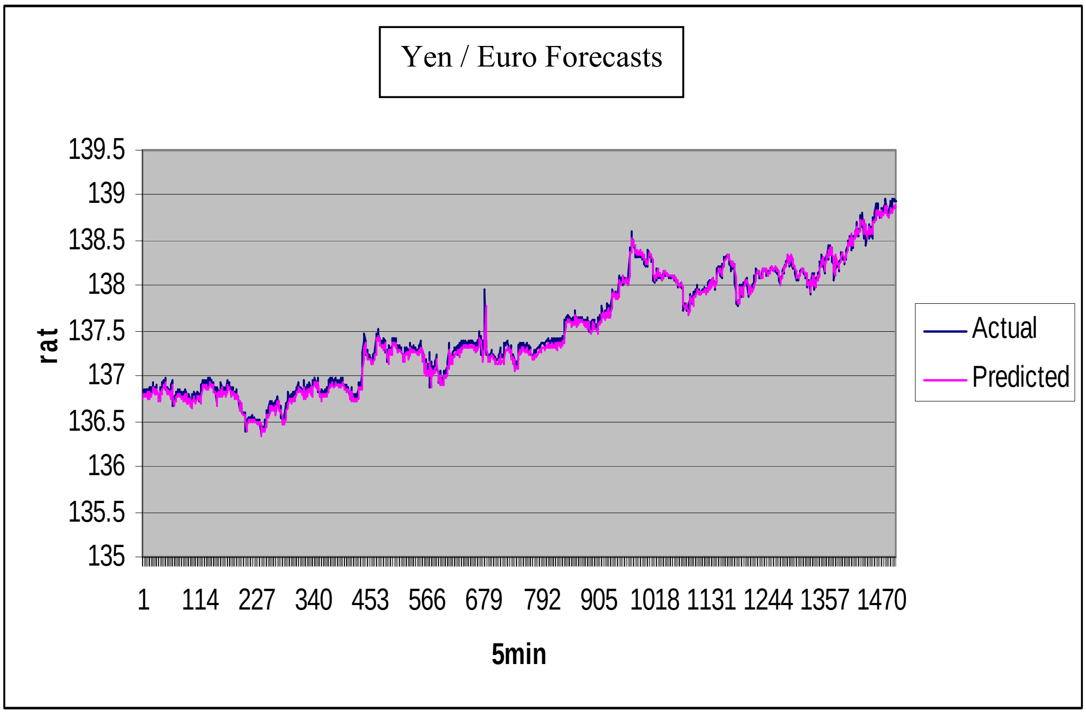

Figure 11: Yen / Euro: Network Architecture 12-5-1 (Testing Phase Data).

5. Conclusions

The reliability of exchange-rate prediction in an international environment of pronounced volatility on one hand and attractive investment opportunities like those arising, for instance, in most of the emerging markets, on the other, has become an issue of outmost importance for both policy makers and investors. Aiming at responding to the challenge of a reliable exchange rate prediction we have used Neural Networks to attain maximum forecasting performance of the U.S. Dollar and the Japanese Yen rates versus the Euro We have started by showing that the R/S analysis can lend a hand to tracing long-term memory in a time series. According to this analysis, the U.S. Dollar / Euro series exhibits no specific long-term pattern while the pattern suggested in the case of the Yen / Euro series is ultimately shown to be misleading, despite initial indications. We have then used the Neuroshell 2 to conclude that Artificial Neural Networks can provide successful time series predictions with a considerable degree of accuracy. We have strong reasons to believe, however, that the use of alternative artificial intelligence methods that can trace the best network topology may improve the prediction accuracy attained in the present paper even further. This requires, of course, further research on the topic that points to the direction of Genetic Algorithms.

Appendix I. R/S Values and Hurst Exponent Estimation.

Step 1 : The original series composed of M observations is modified to one of N = M – 1 observations using logarithmic ratios called returns which are calculated as follows:

$$N_i = \log (M_{(i+1)}/M_i), \text{ where } i = 1, 2, 3\ldots M - 1 \qquad (I.1)$$

The time period spanned by the time series of length N, is divided into m contiguous subperiods of length n such that m*n=N. The elements in each subperiod $X_{i,j}$, have two subscripts, the first (i=1,..,n) to denote the number of elements in each subperiod and the second (j=1,...,m) to denote the subperiod index. For each subperiod j the R/S statistic is calculated, as:

$$\left(\frac{R}{S}\right)_j = s_j^{-1}\left[\max_{1\le k\le n}\sum_{i=1}^{k}(X_{ij} - \overline{X}_j) - \min_{1\le k\le n}\sum_{i=1}^{k}(X_{ij} - \overline{X}_j)\right] \qquad (I.2)$$

where s_j is the standard deviation for each subperiod.

Normalizing (rescaling) the range is important since it permits diverse phenomena and time periods to be compared, which means that R/S analysis can describe time series with no characteristic scale.

Step 2 : The $(R/S)_n$, i.e. the R/S statistic for time length n, is given by the average of the $(R/S)_j$ values for all the m contiguous subperiods with length n, as :

$$\left(\frac{R}{S}\right)_n = \frac{1}{m}\sum_{j=1}^{m}\left(\frac{R}{S}\right)_j \qquad (I.3)$$

Step 3 : Equation (I . 3) gives the R/S value which corresponds to a certain time interval of length n. In order to apply equation (3) as it stands in section 2.1, steps 1 and 2 are repeated by increasing n to the next integer value, until n = N/2, since, at least two subperiods are needed, to avoid bias.

From the above procedure, it becomes obvious that the time dimension is included in the R/S analysis by examining whether the range of the cumulative deviations depends on the length of time used for the measurement. Once equation (I .3) is evaluated for different n periods, the Hurst exponent can be estimated through an ordinary least square regression from equation (3).

Appendix II. Prediction Error Formulas.

$$\mathrm{NRMSE}(n) = \frac{\mathrm{RMSE}(n)}{\sigma_{\Delta}} = \frac{\mathrm{RMSE}(n)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[x_{act}(i) - \overline{x}_{act,n}\right]^{2}}} \qquad \text{(II.1)}$$

where,

$$\mathrm{RMSE}(n) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[x_{pred}(i) - x_{act}(i)\right]^{2}} \qquad \text{(II.2)}$$

$$\mathrm{CC} = \frac{\sum_{i=1}^{n}\left[\left(x_{act}(i) - \overline{x}_{act,n}\right)\left(x_{pred}(i) - \overline{x}_{pred,n}\right)\right]}{\sqrt{\left[\sum_{i=1}^{n}\left(x_{act}(i) - \overline{x}_{act,n}\right)^{2}\right]\left[\sum_{i=1}^{n}\left(x_{pred}(i) - \overline{x}_{pred,n}\right)^{2}\right]}} \qquad \text{(II.3)}$$

$$\mathrm{MRE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_{pred}(i) - x_{act}(i)}{x_{act}(i)}\right| \qquad \text{(II.4)}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|x_{pred}(i) - x_{act}(i)\right| \qquad \text{(II.5)}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(x_{pred}(i) - x_{act}(i)\right)^{2} \qquad \text{(II.6)}$$

where $x_{act}(i)$ and $x_{pred}(i)$ the actual and predicted value when pattern i is presented, $\overline{x}_{act,n}, \overline{x}_{pred,n}$ the mean value of actual and predicted samples of length n and n is the total number of patterns.

References

Amato J, Filardo A, Galati G, Peter G. von and Feng Z, (2005), *Research on Exchange Rates and Monetary Policy: An Overview*, B.I.S. Working Paper No 178.

Andreou, A.S., Georgopoulos, E. and Likothanassis, S. (2002), *Exchange-Rates Forecasting: A Hybrid Algorithm Based on Genetically Optimized Adaptive Neural Networks*, Journal of Computational Economics 20, 3, 191-210.

Andrews, D.: (1991), *Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation,.* Econometrica 59, 817-858.

Anis, A. A. and Lloyd E. H.: (1976), *The Expected Value of the Adjusted Rescaled Hurst Range of Independent Normal Summands,* Biometrica 63, 155-164.

Azoff, E. M. (1994), "Neural Network Time Series Forecasting of Financial Markets", John Wiley & Sons, N.Y.

Baxter, M., (1994), *Real Exchange Rates and Real Interest Differentials,* Journal of Monetary Economics 33, (1), 5-37.

Booth, G. G., F. R. Kaen, and P. E. Koveos, (1982), *R/S Analysis of Foreign Exchange Rates under Two International Monetary Regimes,* Journal of Monetary Economics 10, 407-415.

Brock, W.A. and DeLima, J.F. (1995), "Nonlinear Time Series, Complexity Theory and Finance", The Handbook of Statistics, G.S.Maddala and C.R. Rao, vol. 14, Amsterdam: North Holland.

Campbell, J. and Mankiw, G. (1987), *Are Output Fluctuations Transitory?* Quarterly Journal of Economics 102, 857-880.

Cheung, Y. W. (1993), *Long Memory in Foreign-Exchange Rates,* Journal of Business & Economic Statistics 11, 1, 93-101.

Clark T. and McCraken M. (2001), *Evaluating Long – Horizon Forecasts,* Federal Reserve Bank of Kansas City, Working Paper, 1-14

De Grauwe, P., Dewachter H and Embrechts M, (1993), "Exchange-Rate Theory" Blackwell, Oxford.

Diamandides, P. and Kouretas G. (1996), *Exchange-Rate Determination: Empirical Evidence for the Greek Drachma,* Managerial and Decision Economics 17, 277-290.

Efron, B. (1979), *"Bootstrap Methods: Another Look at the Jack-Knife,* The Annals of Statistics 7, 1-26.

Engel C. and West K. (2004), *Accounting for Exchange-Rate Variability in Present-Value Models when the Discount Factor is Near One*, American Economic Review, Papers and Proceedings, 118 – 125.

Evans M. and Lyons R. (2005), *Meese – Rogoff Redux: Micro-Based Exchange – Rate Forecasting,* N.B.E.R. Working Paper 11042.

Fama, E. and French, K. (1988), *Permanent and Temporary Components of Stock Prices,* Journal of Political Economy 96, 246-273.

Faust J, Rogers J. H. and Wright J. H. (2003), *Exchange Rate Forecasting: The Errors We've Really Made,* Journal of International Economics, 60, 35 – 59.

Fisher, E. (2005), *Exploring Elements of Exchange Rate Theory in a Controlled Envirinment,* Eeconomics.sbs.ohio-state.edu/efisher/exploring3.pdf

Frankel, J. A., (1993), "On Exchange Rates", M.I.T. Press.

Granger, C. and Joyeux, R.(1980), *An Introduction to Long-Range Time Series Models and Fractional Differencing,* Journal of Time Series Analysis 1, 15-30.

Haubrich, J.G. and Lo, A.W. (1989), *The Sources and Nature of Long-Term Memory in Business Cycle*, Working Paper, 5-89, "Rodney L. White" Center for Financial Research, The Wharton School, University of Pennsylvania, Philadelphia.

Haykin, S. (1994), "Neural Networks: A Comprehensive Foundation", MacMillan, New York.

Hiemstra, G. and Jones, J.D. (1995), *Another Look at Long Memory in Common Stock Returns,* Discussion Paper, University of Strathclyde.

Hosking, J. (1981), *Fractional differencing,* Biometrika 68, 165-176.

Hurst, H.E. (1956), *Methods of Using Long-Term Storage in Reservoirs,* Proceedings of the Institute of Civil Engineers 1, 519-543.

Karfakis, C. (1991), *A Model of Exchange-Rate Policy: Evidence for the US Dollar - Greek Drachma Rate 1975 - 1987*, Applied Economics 23, 815-820.

Karytinos, A., Andreou, A.S. and Pavlides, G., (2000), *Long-Term Dependence in Exchange-Rates*, International Journal of Discrete Dynamics in Nature and Society, 4/1, 1-20.

Kilian L. and Taylor M. P. (2003), *Why Is It So Difficult to Beat the Random Walk Forecast of Exchange Rates?* Journal of International Economics, 60, 85 – 107.

Kim, J. C. B. and S. Mo, (1995), *Cointegration and the Long-Run Forecast of Exchange Rates*, Economics Letters 48, 353-359.

Kosko, B. (1992), "Neural Networks for Single Processing", Prentice - Hall.

Koutmos, G. and Theodossiou P. (1994), *Time-Series Properties and Predictability of Greek Exchange Rates*, Managerial and Decision Economics 15, 159-167.

Kuo, C. and Reitsch, A. (1995), *Neural Networks vs. Conventional Methods of Forecasting,* The Journal of Business Forecasting 17-22.

Lewis, K. K., (1989), *Can Learning Affect Exchange-Rate Behavior? The Case of the Dollar in the Early 1980's,* Journal of Monetary Economics 23, 79-100.

Lo, A. W. (1991), *Long-Term Memory in Stock Market Prices,* Econometrica 59, 1279-1313.

Lo, A. and Mackinlay, A.C. (1988), *Stock-Market Prices Do not Follow Random Walks: Evidence From a Simple Specification Test"*, Review of Financial Studies 1, 41-66.

Mandelbrot, B. (1972), *Statistical Methodology for Non-Periodic Cycles: From the Covariance to R/S Analysis,* Annals of Economic and Social Measurement 1.

Mandelbrot, B. and Wallis, J. (1969), *Computer Experiments with Fractional Gaussian Noises*, Parts 1 - 3, Mathematical Appendix. Water Resources Research 5.

Marsh, I. W. and Power D. M. (1996), *A Note on the Performance of Foreign-Exchange Forecasters in a Portfolio Framework*, Journal of Banking & Finance 20, 605-613.

Meese, R. and Rogoff K. (1983), *Empirical Exchange-Rate Models of the Seventies,* Journal of International Economics 14, 3-24.

Mehta, M. (1995), "Foreign-Exchange Markets", in Refenes A.P. (ed.) Neural Networks in the Capital Markets, Wiley, U.K.

Milonas, N., Koveos, P. and Booth, G. (1985), *Memory in Commodity Futures Contracts: A Comment.* The Journal of Futures Markets 5, 113-114.

Noreen, E.W.: (1989), "Computer-Intensive Methods for Testing Hypotheses", New York, John Wiley & Sons.

Patterson, D.W. (1996), "Artificial Neural Networks, Theory and Application", Prentice Hall.

Pesaran, H. M. and Potter S. M. (1992), *Nonlinear Dynamics and Econometrics: An Introduction",* Journal of Applied Econometrics 7, s1-s7.

Peters, E. E. (1994), "Fractal Market Analysis Applying Chaos Theory to Investment and Economics", Wiley Finance Edition.

Pilbeam, K. (1995), *Exchange-Rate Models and Exchange-Rate Expectations: An Empirical Investigation*, Applied Economics 27, 1009-1015.

Pollock, A. C. and Wilkie M. E. (1996), *The Quality of Bank Forecasts: The Dollar-Pound Exchange Rate 1990-1993,* European Journal of Operational Research 91, 306-314.

Poterba, J. and Summers, L.(1988), *Mean Reversion in Stock Returns: Evidence and Implications,* Journal of Financial Economics 22, 27-60.

Refenes, A. P. and Zaidi A. (1995), "Managing Exchange-Rate Prediction Strategies with Neural Networks", in Refenes A.P. (ed.) Neural Networks in the Capital Markets, Wiley, U.K.

Rosenblatt, M. (1956), *A Central-Limit Theorem and a Strong Mixing Condition,* Proceedings of the National Academy of Sciences 42, 43-47.

Rumelhart, D. E. and McLelland J. (1986), "Parallel Distributed Processing", Cambridge, MA, M.I.T. Press.

Steurer, E., (1995), "Non-Linear Modeling of the DEM/USD Exchange Rate", in Refenes A.P. (ed.) Neural Networks in the Capital Markets, Wiley, U.K.

Taylor, M.P. (1995), *The Economics of Exchange Rates,* Journal of Economic Literature 33, 13 – 47

Verrier, M., (1989), "Selection & Application of Currency Forecasts", in C. Dunis & M., Feeny (eds.), Exchange-Rate Forecasting, Probus, Chicago, USA, 303-343.

West, K. D. and D. Cho, (1995), *The Predictive Ability of Several Models of Exchange Rate Volatility",* Journal of Econometrics 69, 367-391.

BANK OF GREECE WORKING PAPERS

21. Kapopoulos, P. and S. Lazaretou, "Does Corporate Ownership Structure Matter for Economic Growth? A Cross - Country Analysis", March 2005.

22. Brissimis, S. N. and T. S. Kosma, "Market Power Innovative Activity and Exchange Rate Pass-Through", April 2005.

23. Christodoulopoulos, T. N. and I. Grigoratou, "Measuring Liquidity in the Greek Government Securities Market", May 2005.

24. Stoubos, G. and I. Tsikripis, "Regional Integration Challenges in South East Europe: Banking Sector Trends", June 2005.

25. Athanasoglou, P. P., S. N. Brissimis and M. D. Delis, “Bank-Specific, Industry-Specific and Macroeconomic Determinants of Bank Profitability”, June 2005.

26. Stournaras, Y., “Aggregate Supply and Demand, the Real Exchange Rate and Oil Price Denomination”, July 2005.

27. Angelopoulou, E., “The Comparative Performance of Q-Type and Dynamic Models of Firm Investment: Empirical Evidence from the UK”, September 2005.

28. Hondroyiannis, G., P.A.V.B. Swamy, G. S. Tavlas and M. Ulan, “Some Further Evidence on Exchange-Rate Volatility and Exports”, October 2005.

29. Papazoglou, C., “Real Exchange Rate Dynamics and Output Contraction under Transition”, November 2005.

30. Christodoulakis, G. A. and E. M. Mamatzakis, “The European Union GDP Forecast Rationality under Asymmetric Preferences”, December 2005.

31. Hall, S. G. and G. Hondroyiannis, “Measuring the Correlation of Shocks between the EU-15 and the New Member Countries”, January 2006.

32. Christodoulakis, G. A. and S. E. Satchell, “Exact Elliptical Distributions for Models of Conditionally Random Financial Volatility”, January 2006.

33. Gibson, H. D., N. T. Tsaveas and T. Vlassopoulos, “Capital Flows, Capital Account Liberalisation and the Mediterranean Countries”, February 2006.

34. Tavlas, G. S. and P. A. V. B. Swamy, “The New Keynesian Phillips Curve and Inflation Expectations: Re-specification and Interpretation”, March 2006.

35. Brissimis, S. N. and N. S. Magginas, "Monetary Policy Rules under Heterogeneous Inflation Expectations", March 2006.

36. Kalfaoglou, F. and A. Sarris, "Modeling the Components of Market Discipline", April 2006.

37. Kapopoulos, P. and S. Lazaretou, "Corporate Ownership Structure and Firm Performance: Evidence from Greek Firms", April 2006.

38. Brissimis, S. N. and N. S. Magginas, "Inflation Forecasts and the New Keynesian Phillips Curve", May 2006.

39. Issing, O., "Europe's Hard Fix: The Euro Area", including comments by Mario I. Blejer and Leslie Lipschitz, May 2006.

40. Arndt, S. W., "Regional Currency Arrangements in North America", including comments by Steve Kamin and Pierre L. Siklos, May 2006.

41. Genberg, H., "Exchange-Rate Arrangements and Financial Integration in East Asia: On a Collision Course?", including comments by James A. Dorn and Eiji Ogawa, May 2006.

42. Christl, J., "Regional Currency Arrangements: Insights from Europe", including comments by Lars Jonung and the concluding remarks and main findings of the workshop by Eduard Hochreiter and George Tavlas, June 2006.

43. Edwards, S., "Monetary Unions, External Shocks and Economic Performance: A Latin American Perspective", including comments by Enrique Alberola, June 2006.

44. Cooper, R. N., "Proposal for a Common Currency Among Rich Democracies" and Bordo, M. and H. James, "One World Money, Then and Now", including comments by Sergio L. Schmukler, June 2006.

45. Williamson, J., "A Worldwide System of Reference Rates", including comments by Marc Flandreau, August 2006.

46. Brissimis, S. N., M. D. Delis and E. G. Tsionas, "Technical and Allocative Efficiency in European Banking", September 2006.

47. Athanasoglou, P. P., M. D. Delis and C. K. Staikouras, "Determinants of Bank Profitability in the South Eastern European Region", September 2006.

48. Petroulas, P., "The Effect of the Euro on Foreign Direct Investment", October 2006.