*Discrete Dynamics in Nature and Society*, Vol. 5, pp. 121–137
Reprints available directly from the publisher
Photocopying permitted by license only



# In Search of a Warning Strategy Against Exchange-rate Attacks: Forecasting Tactics Using Artificial Neural Networks

A. S. ANDREOU[a,*], G. A. ZOMBANAKIS[b,†], E. F. GEORGOPOULOS[c,‡] and S. D. LIKOTHANASSIS[c,¶]

[a]*University of Patras, Department of Computer Engineering and Informatics, and Artificial Intelligence Research Center (U.P.A.I.R.C.);* [b]*Bank of Greece, Research Department;* [c]*University of Patras, Department of Computer Engineering and Informatics, Artificial Intelligence Research Center (U.P.A.I.R.C.) and Computer Technology Institute*

*(Received 30 December 1999)*

"**Heart attacks and devaluations are not predictable and, certainly, are never preannounced**". (The usual remark made by government spokesmen shortly after a domestic currency devaluation has taken place.)

The contribution that this paper aspires to make is the prediction of an oncoming attack against the domestic currency, something that is expected to increase the possibilities of successful hedging by the authorities. The analysis has focused on the Greek Drachma, which has suffered a series of attacks during the past few years, thus offering a variety of such "shock" incidents accompanied by frequent interventions by the authorities. The prediction exercised here is performed in a discrete dynamics environment, based on the daily fluctuations of the interbank overnight interest rate, using artificial neural networks enhanced by genetic algorithms. The results obtained on the basis of the forecasting performance have been considered most encouraging, in providing a successful prediction of an oncoming attack against the domestic currency.

*Keywords*: Shocks; Interest-rate; Forecasting; Neural-networks; Genetic algorithms

## 1. INTRODUCTION

Exchange-rate crises represent a menace all central banks have learned to live with, fight against and suffer its consequences. It is only natural, therefore, that any contribution to the arsenal of crises-fighting devices used by the authorities must be more than welcome. This paper aims at

*Tel.: +061 997755, Fax: +061 997706, e-mail: aandreou@ceid.upatras.gr
†Address for correspondence: Bank of Greece, Research Department, 21 Panepistimiou str., Athens 10250, Greece. Tel.: +301 3235809, Fax: +01 3233025, e-mail: boglibr@ath.forthnet.gr
‡e-mail: georgops@ceid.upatras.gr
¶e-mail: likothan@cti.gr

contributing to crises management by proposing an application of selected computational intelligent schemes based on artificial neural networks for the prediction of an oncoming currency shock in an environment of discrete dynamics of the time-series data involved.

The currency we have decided to study is the Greek drachma during the period between January, 1990 and February, 1999, a choice based on the following reasons: First, the drachma has been the victim of a considerable number of speculative attacks, which have intensified during the last few years, and led to its devaluation, despite the "hard-drachma" policy aimed at curbing inflation. Second, the Greek economy has been subjected to a wide variety of structural and institutional reforms during the period under consideration, which have encouraged these speculative attacks. Finally, Greece has gone through a number of pre-election periods, during which the drachma was put under severe pressure as a result of the market expectations.

The reaction of the Central Bank in attempting to face attacks against the drachma involves, in most cases, the use of two policy instruments: The daily reserves spent during a market intervention and the overnight interest rate ($o/n$) which is the daily borrowing rate among commercial banks or between the Central Bank and commercial banks. Data on daily interventions by the Central Bank are, unfortunately, not available, due to international commitments of the local authorities. This leaves us with just one instrument, the overnight interest rate, which is raised in order to protect the domestic currency rates from the adverse effects of exogenous disturbances, like the ones arising when the local currency becomes the target of a speculative attack. In such cases, raising the overnight rate acts as a bumper against the attack and leaves the local currency exchange rates unaffected. It is important to understand, therefore, that a study of these rates themselves is not expected to reveal the effect of an attack, since the full effect of such an attack is neutralized due to the market intervention by the Central Bank. It is equally important, in addition, to agree on defining a period of "crisis" or "shock" as one during which the interbank overnight interest rate rises in protection of the domestic currency rates. This rise must be both substantial (in percentage terms with respect to the pre-crisis rates), as well as sustained in terms of duration.

A thorough study of the drachma exchange-rate history reveals that the Greek currency has been subjected to attacks of varying character, intensity and duration, as reflected in the fluctuations of the overnight interest rate, depending on whether the cause that triggered the particular shock is associated with a pre-election period, a major domestic market reform or an unrest in the international markets, like the ones that took place in the Southeast Asia, Russia or Brazil. We shall be able to see during the analysis, that the forecasting method employed in this paper is based on training the algorithm to recognize the particular pattern of overnight-rates fluctuations which represents the reactions of the authorities during periods of shocks like the ones described above, learning to expect a similar behavior in the future, once this particular pattern of data fluctuations is again encountered. The "reasoning" used by this method leads to conclusions which should by no means be considered as restricted only to the drachma case, since the networks used are trained to recognize specific patterns of data behavior during the training period, irrespective of the nature of the variable itself. Besides, the expected EMU membership of Greece beginning in the year 2001 will eventually relieve the drachma from the imminent danger of such an attack. One may safely argue, therefore, that the results derived in this paper may be considered as holding true in the case of most small, open economy currencies.

The paper is organized as follows: Section 2 involves a literature overview, while Section 3 describes the economic environment in Greece. The methodology used, as well as the technical framework on which the paper is based, are analyzed in Section 4, while the empirical findings

and the conclusions drawn are presented in Sections 5 and 6 respectively.

## 2. LITERATURE OVERVIEW

Predictability issues referring to foreign exchange markets, have led, in most cases, to results which are either limited (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). The difficulties encountered are attributed, among other things, to the noise level introduced by frequent policy interventions from the part of the authorities (Taylor, 1995), and to the failure of the structural models to outforecast the random walk model, due, among other things, to difficulties in modelling expectations of the explanatory variables (Meese and Rogoff, 1983; Leventakis, 1987; De Grauwe *et al.*, 1993; Frankel, 1993; Baxter, 1994 and Pilbeam, 1995). The empirical failure of models to forecast exchange-rate movements may also be due to the fact 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 patterns bring about significant influence on the market movements (Mehta, 1995).

Thus, several authors seem to conclude that even the forward rate, which is considered very efficient when used to improve forecasting performance, can sometimes fail in contributing towards this direction (Levich, 1989). These difficulties in predicting the exchange rate are accentuated when it comes to the question of currency-crises predictability (Berg and Pattillo, 1999), and even more so in the case of a devaluation, since, apart from the issue of technical problems, one has to face the fact that the decision to devalue is primarily a policy-maker's option, which is not at all certain to follow an attack against the domestic currency. In a number of cases, however, some authors deal with the matter in terms of forecasting the rate of devaluation in the ERM context (Rose and Svenson, 1995), while others (Siklos and Tarajos, 1996) try to determine the probability and expected devaluation rate on the basis of monthly cross-section data. Koedijk and Kool (1994) deal with speculative strategies within the EMS depending on the ability to predict the timing and the rate of the parity change, while an interesting case study by Majuca (1992) attempts to predict the date of the exchange-rate collapse of the Philippinese peso.

As has been admitted in the literature, the limited success in interpreting exchange-rate movements, has led to the use of "some recently developed sophisticated time-series techniques" (Taylor, 1995), tracing chaotic behaviour in the exchange-rates series examined, as well as the method of artificial neural networks. These methods, being data-driven approaches, have been considered preferable to traditional, model-driven approaches used for forecasting purposes. In fact, the exchange-rate literature has been recently enriched by an increasing number of studies which resort to using the neural networks methodology for exchange-rate forecasting and lead to better results compared with "conventional methods" (*e.g.*, Mehta, 1995; Steurer, 1995; Refenes and Zaidi, 1995). With reference to Greece and the drachma rates, in particular, Karytinos *et al.* (1999) and Andreou *et al.* (2000) have focused on revealing long-term dependence and the extent to which the time series involved exhibit chaotic behaviour, while another series of papers attempt to predict the exchange rate of certain major currencies against the Greek Drachma (Andreou *et al.*, 1997 and 1998a; Adamopoulos *et al.*, 1997). The case of classifying a period as normal or under attack using feedforward Multi-Layer Perceptrons (MLP's) neural networks is reported in Andreou *et al.* (1998b) yielding successful results.

## 3. ECONOMIC ENVIRONMENT AND SHOCKS

The ineffectiveness of the "fully-accommodating" exchange-rate policy of the 80's, leading to the depreciation-inflation vicious cycle by generating inflationary expectations is a fact extensively underlined by various authors like Brissimis and Leventakis (1989); Karadeloglou (1990); Zombanakis (1998) and Karadeloglou *et al.* (1998) and by the authorities themselves (Bank of Greece 1991, p. 33). With the business sector neglecting any supply-side reforms while relying heavily on the state support expressed as a subsidy via the drachma depreciation, the policy makers had no choice but to resort to a non-accommodating depreciation policy since the early '90s, and later on, to fixing the exchange rate of the drachma vis-à-vis the ECU. This policy has proven to be more than successful, since its drastic anti-inflationary impact was accompanied by a significant interest-rate reduction representing a relief for the budget deficit and a decrease of the capital cost of the business sector. Thanks, also, to the "hard-drachma" policy, increases in servicing the external debt have been avoided and the foreign-exchange risk has been restricted. Meanwhile, the cost of the import component for Greek export firms has been held constant, in view of attaining the economy's international, long-run policy targets (Bank of Greece, 1994) and catching "the last train to Euro" in January, 2001.

The persistent and intensive pressure on the Greek currency following the Southeast Asia crisis, however, led to its ERM participation, along with its devaluation on March 13, 1998, despite the adverse impact on the number one target, *i.e.*, the inflation rate (Bank of Greece, 1997). This prolonged run on the drachma, with investors selling Greek bonds to make up for the losses suffered in the Southeast Asian markets, lasted about a month, sometime between October and November, 1997. The overnight rate, nevertheless, rose by about 950%, a percentage indicative of the intensity of the crisis. This was succeeded by a prolonged period of continuous episodes, persistent in character, although moderate in intensity. Thus, the pressure on the drachma continuing throughout the beginning of 1998 in the form of a series of brief, repeated attacks, each of small duration, leads to the conclusion that the measures taken to face this run on the drachma have proven inadequate to avert its devaluation. We believe, therefore, that this behavioural syndrome composed of a sequence of two or more phases of a crisis, a major, however relatively brief attack followed by a second (or even several others), in an environment of specific exchange-rate policy commitments, should be regarded as an alarming indication which the market cannot afford to ignore. A word of caution is required at this point, however: This statement does not imply that all such behavioural complexes necessarily lead to devaluations. It simply asks the analyst to consider it as being unusual and, therefore, worth noting, given the particular exchange-rate policy pursued by the authorities. After all, only one of two such major runs during the period under review, namely the one just described, did indeed lead to a devaluation. The second, *i.e.*, the one of 1994 caused as a reaction of the market to institutional reforms, was much less painful, as we shall see later on.

A closer look at the remaining most important incidents in the 1990s which have resulted in "shocks" against the Greek currency, leads to the conclusion that politics play indeed a major role. In fact, the period 1989 to 1990 offers a good example of a political shock of the domestic currency due to devaluation expectations in the market in the context of a prolonged pre-election period. The impact of the October, 1993 parliamentary elections on the domestic market was also pronounced, albeit to a lesser extent compared to that of 1990, with the overnight rate increasing by about 45%, and its duration being as long as four to five months.

The period under review here is full of structural reforms for the Greek economy, with the May, 1994 removal of the capital movement restrictions

TABLE I Main statistics and special features of selected runs against the Greek Drachma (1990–1998)

| | 1989/1990 Election | 1992 ERM | 1993 Election | 1994 Policy | 1997 S.E.Asia | 1998 Deval. | 1998 Russia |
|---|---|---|---|---|---|---|---|
| Av. Dev. | 5.24 | 4.07 | 2.5 | 45.09 | 22.72 | 2.47 | 1.03 |
| Var. Pop. | 36.31 | 23.15 | 8.55 | 2582.98 | 1139.26 | 11.26 | 1.63 |
| Dev. Sq. | 8422.7 | 1134.5 | 1248.5 | 51659.6 | 37595.5 | 1013.5 | 43.9 |
| Coef. Var. | 0.288 | 0.143 | 0.120 | 0.489 | 1.050 | 0.249 | 0.101 |
| Duration | 7.5 months | 1.5 month | 4.5 months | 20 days | 1 month | 3 months | 1 month |
| Max. $o/n$ Change | 190% | 70% | 45% | 860% | 950% | 170% | 40% |

mentioned above being of major importance. Since this "institutional" reform had been more or less anticipated by the market, the overnight interest rate reached a global maximum of about 860% during the second week of that "crisis". The successful treatment of this attack on the basis of the overnight rate was supplemented by the outflow of about $3 billion of the Central Bank's foreign exchange reserves to maintain the exchange rate of the drachma and drove the crisis to an end after a brief duration of about two weeks.

Furthermore, the international crisis that started on the 16th of September, 1992 affected the Greek market rather moderately, since the Central Bank had to raise the overnight rate by about 70%, just for one day, a score and a duration which are rather low compared to those of other cases. The consequences of the turmoil in the international markets, however, kept affecting the domestic economy for a month or so. Finally, the impact of the August, 1998 Russian crisis on the Greek economy seems similar to some extent, to that caused in September, 1992, despite the differences in the nature and causes between the two incidents, with the pressure on the drachma lasting for about a month and the overnight rate rising by roughly 40%.

A comprehensive picture of these events in terms of their main statistics is included in Table I.

## 4. TECHNICAL BACKGROUND

This section is devoted to introducing and analyzing the technique of artificial neural networks, which belongs to a class of data-driven approaches, as opposed to model-driven approaches. Certain general-purpose algorithms address the process of constructing such a "machine" based on available data. The problem is then reduced to the computation of the weights of a feedforward network to accomplish a desired input–output mapping and can be viewed as a high-dimensional, nonlinear system identification problem of discrete dynamics. In a feedforward network, the units can be partitioned into layers, with links from each unit in the $k$th layer being directed to each unit in the $(k+1)$th layer. Inputs from the environment enter the first layer and outputs from the network are manifested in the last layer. An $m$-$d$-1 architecture, shown in Figure 1, refers to a feedforward Multi-Layer Perceptron (MLP) network with $m$ inputs, $d$ nodes in the hidden layer and one node in the output layer.

From the given time series $x = \{x(t)\colon 1 \le t \le N\}$ of exchange-rate data, we obtain two sets: a training set $x_{\text{train}} = \{x(t)\colon 1 \le t \le T\}$, and a test set $x_{\text{test}} = \{x(t)\colon (T+1) \le t \le N\}$, where $N$ is the length of the data record. After a period of learning, during which the training patterns and corresponding actual values are presented iteratively to the network, convergence is achieved. The evaluation of the network's performance is carried out using the testing set, a task that investigates whether the network succeeded in generalizing the knowledge embodied through the training phase rather than simply memorising it.

The forecasting task was based on two different techniques from the implementation point of view:

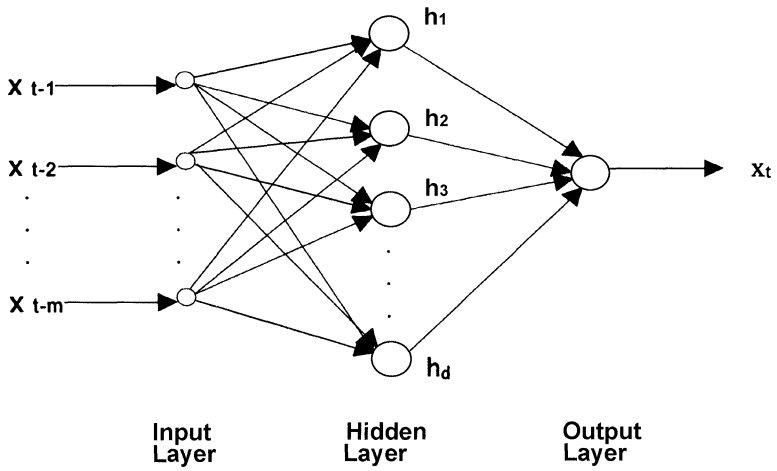



FIGURE 1 The *m-d*-1 feedforward Multi-Layer Perceptron (MLP) neural network architecture, where *m* denotes the number of inputs and *d* the number of hidden nodes.

The first approach employs MLP networks trained by an algorithm relying on a Kalman filter for training and a genetic algorithm for evolving the structure of nodes in the input and hidden layers. The genetic algorithm monitors the course of forecasting and alters the size of the network to obtain the optimum architecture. The Kalman filter was selected for training due to its speed, during the recurrent process of learning which is provided by its following two features (Haykin, 1994): (i) Efficient utilization of the information contained in the input data using state-space concepts, and (ii) Estimation of the state recursively from the previous estimate and the currently available data, thus reducing storage requirements.

The second approach uses a flexible multiple layer perceptron scheme, in which each layer is activated by a different function. Thus, the task of adjusting the internal weights of the network relies on multiple non-linear computations, which leads to better fine-tuning of the internal parameters. This second scheme will provide us with the opportunity to incorporate an additional variable as input to the networks, which measures the intensity of a shock period, as it will be described below.

### 4.1. The Localized Extended Kalman Filter

A category of feed-forward ANN (Artificial Neural Network) training algorithms are those using advanced filtering techniques. One of the most effective among the different methods belonging to this category, is the localized approach of the Extended Kalman Filter, briefly presented in this section.

Let us consider a network characterized by a weight vector $w$ representing the free parameters of the network. The average cost function that should be minimized during the training phase is defined in terms of $N$ input–output patterns as follows:

$$E_{av}(w) = \frac{1}{2N} \sum_{n=1}^{N} \sum_{j \in C} \left[ d_j(n) - y_j(n) \right]^2 \qquad (1)$$

where $d_j(n)$ is the desired response and $y_j(n)$ the actual response of output neuron $j$ when input pattern $n$ is presented, while the set $C$ includes all the output neurons of the network. The cost function $E_{av}(w)$ depends on the weight vector $w$ due to the fact that $y_j(n)$ itself depends on $w$.

Concentrating on an arbitrary neuron $i$, which might be located anywhere in the network, its behavior during the training phase may be viewed

as a nonlinear dynamic system, which, in the context of Kalman filter theory, may be described by the following state-measurement equations (Haykin, 1994; Shah *et al.*, 1992):

$$w_i(n+1) = w_i(n) \tag{2}$$

$$d_i(n) = y_i(n) + e_i(n) \tag{3}$$

$$y_i(n) = \varphi(x_i^T(n), w_i(n)) \tag{4}$$

Here, the iteration $n$ corresponds to the presentation of the $n$th input pattern, $x_i(n)$ and $y_i(n)$ are the input and output vector of neuron $i$ respectively and $e_i(n)$ is the measurement error at the output of neuron $i$, the instantaneous estimate of which is given by:

$$e_i(n) = -\frac{\partial E(n)}{\partial y_i(n)} \tag{5}$$

$$E(n) = \frac{1}{2}\sum_{j \in C}[d_j(n) - y_j(n)]^2 \tag{6}$$

The differentiation in Eq. (5) corresponds to the back-propagation of the global error to the output of neuron $i$.

The activation function $\varphi(\cdots)$ is responsible for the nonlinearity in the neuron. The weight vector $w_i$ of the optimum model for neuron $i$ is to be "estimated" through training with examples. The activation function is assumed to be differentiable. Accordingly, we can use Taylor series to expand Eq. (3) about the current estimate of the weight vector and thereby linearize the equation as follows (Haykin, 1994):

$$\varphi(x_i^T(n)w_i(n)) \cong q_i^T(n)w_i(n) + [\varphi(x_i^T(n)\hat{w}_i(n)) - q_i^T(n)\hat{w}_i(n)] \tag{7}$$

where,

$$q_i(n) = \left[\frac{\partial \varphi(x_i^T(n)w_i(n))}{\partial w_i(n)}\right]_{w_i(n)=\hat{w}_i(n)} = \hat{y}_i(n)[1 - \hat{y}_i(n)]x_i(n) \tag{8}$$

$\hat{y}_i(n)$ is the output of neuron $i$ that results from the use of the weight estimate. In Eq. (7) we have assumed the use of the logistic function:

$$\left[1 + \exp\left(-\sum_{i=1}^{m} w_i x_i + \theta\right)\right]^{-1} \tag{9}$$

Other sigmoid functions, like the hyperbolic tangent, can be used as well. The first term of the right hand side of Eq. (7) is the desired linear term while the remaining term represents a modeling error. Thus, substituting Eqs. (7) and (4) in (3) and ignoring the modeling error we obtain:

$$d_i(n) = q_i^T(n)w_i(n) + e_i(n) \tag{10}$$

where $e_i(n)$ and $q_i(n)$ are defined in Eqs. (5) and (8) respectively. Equations (2) and (10) describe the linearized behavior of neuron $i$.

Given the pair of Eqs. (2) and (10), we can make use of the standard Recursive Least Squares (RLS) algorithm equations (Haykin, 1994), which is a special case of the Kalman filter, to make an estimate of the weight vector $w_i(n)$ of neuron $i$. The resulting solution is defined by the following system of recursive equations (Haykin, 1994), that describe the Multiple Extended Kalman Algorithm (MEKA) (Shah *et al.*, 1992):

$$r_i(n) = \lambda^{-1}P_i(n-1)q_i(n) \tag{11}$$

$$k_i(n) = r_i(n)[1 + r_i^T(n)q_i(n)]^{-1} \tag{12}$$

$$w_i(n+1) = w_i(n) + e_i(n)k_i(n) \tag{13}$$

$$P_i(n+1) = \lambda^{-1}P_i(n) - k_i(n)r_i^T(n) \tag{14}$$

where, $n = 1, \ldots, N$ is the iteration number and $N$ is the total number of examples. The vector $q_i(n)$ represents the linearized neuron activation function given in Eq. (8), $P_i(n)$ is the current estimate of the inverse of the covariance matrix of $q_i(n)$ *and* $k_i(n)$ *is the Kalman gain*. The parameter $\lambda$ is a

*forgetting factor* which takes values in the range (0,1], and $e_i(n)$ is the localized measure of the global error. Equation (14) is called the Riccati difference equation (Haykin, 1994). Each neuron in the network perceives its own effective input $q_i(n)$, hence it has to maintain its own copy of $P_i(n)$, even in the case in which it may share some of its inputs with other neurons in the network.

### 4.2. The Hybrid Algorithm

MLP neural network architectures are evolved through a genetic algorithm, each being trained by MEKA to produce predictions on the overnight interest rate. In particular, a proposed modified Genetic Algorithm (GA) maintains a population of individuals (Neural Networks) for each generation, having random structure in the hidden region. The MEKA algorithm is employed for the training of each network for just one epoch. Performance is measured with the fitness function, which is a function of the MRE described later on and the size (number of nodes) of the network. Then a new population is created, by selecting the top individuals based on their fitness (select step). Some members of the population undergo transformations by means of genetic operators to form the new individuals. We use a mutation operator that changes the structure of the network randomly in order to preserve diversity. Also, there is a crossover operator, which creates new individuals by combining parts from two individuals. After some number of iterations the program converges at a near-optimum solution.

The steps of the genetic algorithm are analytically described as follows:

*Step 1*, *Initialization* An initial population of randomly generated individuals (random number of inputs and hidden neurons) is created. Generally, a large population size is preferable, but in our experiments we need to compromise with the computer limitations, so a population of fifty individuals was used in all of the conducted experiments. The connection weights are initialized to random values in $[-1,1]$, using uniform probability distribution.

*Step 2*, *Selection* Selection is an essential operation in genetic algorithms; it constructs a new population with respect to the probability distribution based on fitness values of the individuals of the previous population. In our experiments, a variation of the classic Roulette Wheel Selection Operator (Michalewicz, 1996) was used. In this variation we save the best ever individual in a place outside the population and in the selection operation we make sure that at least one copy of this individual will pass to the next generation (elitism).

The fitness function used in the selection phase takes into account the performance of the network on the test set and its size and has the following form:

$$\text{Fitness} = 1/(1 + \text{MRE} + \text{size\_par} * \text{MRE} * \text{SIZE}) \tag{15}$$

where size_par is a parameter that controls the importance of a network's size in the evaluation of the fitness function. The objective for size_par is to take values that will lead to individuals with small sizes, maintaining though good forecasting ability. The term (size_par*MRE*SIZE) allows for the importance of the network's size to decrease accordingly to the decrease of MRE.

*Step 3*, *Crossover* The crossover operator is applied to the new population. Generally it works as follows: it selects two parents and generates one or two offsprings by recombining parts of them. The offsprings take the place of their parents in the new population. In the proposed algorithm crossover operates as follows:

Let us assume that we have the two parents: $I_1H_1O$ and $I_2H_2O$ where $I$, $H$ and $O$ are the numbers of input, hidden and output nodes, respectively. Next we generate the random numbers:

$$i_j = \text{a uniform random number in } [0, I_j], \quad j = 1,2$$
$$h_j = \text{a uniform random number in } [0, H_j], \quad j = 1,2.$$

Then we create a child with $(i_1+i_2)$ input notes, $(h_1+h_2)$ hidden nodes and $O$ output nodes. If $(i_1+i_2)=0$ then we set the number of input nodes to 1; if $(h_1+h_2)=0$ we set the number of hidden nodes to 1. The weights of the child are initialized randomly in the same interval that was used in the initialization phase. The second child is created in the same manner.

*Step 4, Mutation* The mutation operator works as follows: It selects at random a neural network (individual) from the population and changes its number of inputs and/or its number of hidden neurons by adding or deleting a random number (selected uniformly from a given interval) of inputs and/or hidden neurons.

### 4.3. The Multi-layer Multiply Activated Perceptron (MLMAP)

The basic component of this technique is a multilayer perceptron as shown in Figure 2. The network consists of three hidden layers connected to each other. Each layer has its own activation function:

The first hidden layer uses the hyperbolic tangent given in:

$$\tanh(x) = \frac{1 - e^{-bx}}{1 + e^{-bx}}. \tag{16}$$

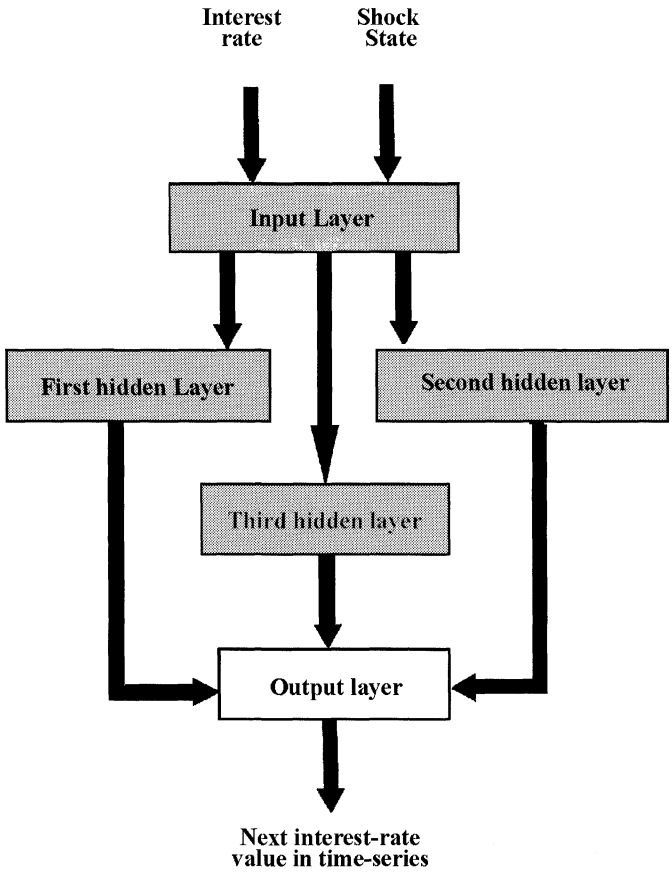



FIGURE 2 Graphical representation of the layers structure in the Multi-Layer Multiply Activated Perceptron (MLMAP) neural network scheme.

The second one uses the Gaussian as its activation function:

$$f(x) = e^{-x^2}. \tag{17}$$

Finally, the third hidden layer is based on the Gaussian complement activation function:

$$f(x) = 1 - e^{-x^2}. \tag{18}$$

The input layer is linear, while the output uses the sigmoid function.

### 4.4. Performance Evaluation

Evaluation of forecasting is performed through three well-known error measures, the Normalized Root Mean Squared Error (NRMSE), the Correlation Coefficient (CC) and the Mean Relative Error (MRE). All these measures were applied on the testing set of data, that is, a set of patterns that did not participate during the course of learning.

The Normalized Root Mean Square Error (NRMSE) is defined by:

$$\begin{aligned} \mathrm{NRMSE}(n) &= \frac{\mathrm{RMSE}(n)}{\sigma_\Delta} \\ &= \frac{\mathrm{RMSE}(n)}{\sqrt{(1/n)\sum_{i=1}^{n}[x_{\mathrm{act}}(i) - \bar{x}_n]^2}} \end{aligned} \tag{19}$$

where,

$$\mathrm{RMSE}(n) = \sqrt{(1/n)\sum_{i=1}^{n}[x_{\mathrm{pred}}(i) - x_{\mathrm{act}}(i)]^2} \tag{20}$$

On the other hand, the correlation coefficient (CC) between the actual and predicted series is given by

$$\mathrm{CC} = \frac{\sum_{i=1}^{n}[(x_{\mathrm{act}}(i) - \bar{x}_{\mathrm{act,n}})(x_{\mathrm{pred}}(i) - \bar{x}_{\mathrm{pred}}, n)]}{\sqrt{\left[\sum_{i=1}^{n}(x_{\mathrm{act}}(i) - \bar{x}_{\mathrm{act,n}})^2\right]\left[\sum_{i=1}^{n}(x_{\mathrm{pred}}(i) - \bar{x}_{\mathrm{pred,n}})^2\right]}}. \tag{21}$$

The CC measures the ability of the predicted samples to follow the upward or downward jumps of 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 with a negative mirroring, that is with a 180° rotation about the time-axis. When the original series moves up, the forecasting moves down at the same time-period and *vice versa*.

Finally, the MRE is given by the formula:

$$\mathrm{MRE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_{\mathrm{pred}}(i) - x_{\mathrm{act}}(i)}{x_{\mathrm{act}}(i)}\right| \tag{22}$$

where $x_{\mathrm{pred}}(i)$ is the output of the network, $x_{\mathrm{act}}(i)$ is the actual value when pattern $i$ is presented and $n$ is the total number of patterns. MRE shows the percentage of accuracy of predictions expressing it in a stricter way, since it focuses on the sample being predicted. Thus, we are able to estimate prediction error as a fraction of the actual value.

## 5. EMPIRICAL RESULTS

### 5.1. Application of MEKA and MLMAP

As has already been indicated, the two major shocks suffered during the period under consideration are one in May 1994, as a result of the complete liberalisation of capital movements by the removal of all institutional barriers imposed on international capital transactions, and the second one during and after the crisis in the Southeast Asia, during the second half of 1997 (Fig. 3). The training stage, therefore, provides the algorithms used with the pattern of the overnight-rate fluctuations exhibited as a result of the reaction of the Central Bank to support the drachma during these two crises, as well as during several milder ones, mostly caused by political rather than economic reasons. The two computational intelligent systems employed, namely the MEKA and the MLMAP, have been subjected to repeated experiments with a forecasting horizon of one and

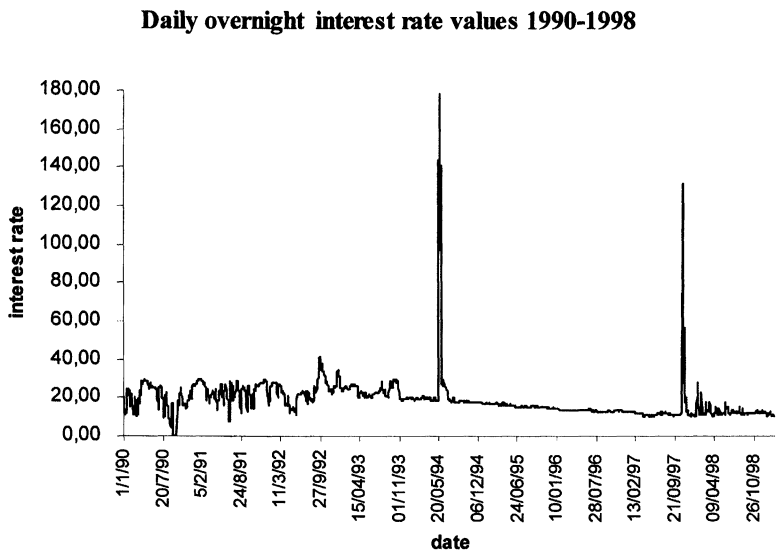



FIGURE 3 Daily values for the Greek overnight interest rate. Sharp peaks correspond to periods of attacks against the drachma.

five days in each case. To make up for the lack of information on the daily foreign-exchange reserves spent by the Central Bank to support the drachma during periods of crises, we have made use of the MLMAP technique which allows for the introduction of a dummy variable operating as a shock flag or indicator. This is a triple-state variable encoding the effects of the shock period it indicates: Thus, a value of 1 indicates a period characterized by severe pressure on the drachma, leading to its devaluation, while a value of 0.5 denotes an attack against the drachma which, was successfully faced by the authorities without any repercussions on the drachma exchange rate. Finally the value of 0 corresponds to a perfectly normal period, free of attacks of any kind.

Table I indicates that the two major attacks mentioned above, have indeed led to spectacular increases of the *o/n* rates of the order of 900%, in defence of the drachma parity. In the fist case, in May 1994, this rise, accompanied by an unknown, nevertheless substantial amount of foreign-reserves outflow, was enough to handle the crisis that lasted only about 20 days. The second run started with the crisis in Southeast Asia during the second half of 1997, and, despite the pronounced reaction of the *o/n* rate and the fact that the Central Bank spent, until the end of the year, at least 10 billion dollars of foreign exchange reserves to protect the drachma, it led to the March 1998 devaluation.

The results reported below were derived on the basis of a 1200-sample length for the training data set (approximately 5 years, in-sample data) covering the period from 1/1/90 to 30/6/97 and a 580-sample length (approximately 2,5 years out-of-sample data) for the testing set, from 1/7/97 to 10/2/99 starting immediately after the last sample value of the training set. Objective evaluation on this testing set of data, which was not previously used during any of the intermediate networks training stages is achieved through the error measures reported earlier. The set of inputs used for the MEKA method is decided by the genetic algorithm evolving the structure of the networks. The MLMAP input set consists of the daily *o/n* interest rate values during the period mentioned above and the shock flag, while the number of hidden nodes have been empirically chosen through numerous runs.

Table II summarises the best results produced for each of the two methods applied on a daily and

TABLE II Evaluation of interest-rate prediction during the training and testing phases for a one and five-days prediction horizon

| Horizon | Network architecture* | Training phase | | | Testing phase | | |
|---|---|---|---|---|---|---|---|
| | | C.C. | NRMSE | MRE | C.C. | NRMSE | MRE |
| 1 | MEKA 4-5-1 | 0.9640 | 0.3443 | 0.0610 | 0.7503 | 0.7517 | 0.0629 |
| 5 | MEKA 5-17-1 | 0.7674 | 0.6735 | 0.1133 | 0.2278 | 1.8401 | 0.2840 |
| 1 | MLMAP 2-10-10-18-1 | 0.967 | 0.253 | 0.033 | 0.728 | 0.814 | 0.168 |
| 5 | MLMAP 2-10-10-18-1 | 0.794 | 0.610 | 0.086 | 0.245 | 0.978 | 0.130 |

*MEKA $p$-$k$-$t$ refers to MEKA algorithm on a $p$-inputs, $k$-hidden nodes and $t$ outputs network MLMAP $p$-$k_1$-$k_2$-$k_3$-$t$ refers to MLMAP scheme with $p$-inputs, $k_1$ nodes in the first hidden layer, $k_2$ nodes in the second hidden layer, $k_3$ nodes in the third hidden layer and $t$ outputs.

five-days prediction horizon. The MEKA algorithm used a number of generations equal to 1000, a population size equal to 50, while the *size_par* variable was set to $10^{-4}$. The results are quite encouraging. The generalisation ability of all networks is satisfactory, particularly when it comes to the one-day prediction horizon, yielding more accurate predictions, as expected. Daily forecasts produced by MEKA are slightly superior to those of the MLMAP, reaching a success level of 75% for the former and 73% for the latter in CC (Correlation Coefficient) terms during the testing phase. The NRMSE measure (Normalised Root Mean Square Error) reveals a high predictive ability for both schemes, clearly superior to that of a simple mean predictor in both horizons, with MEKA being superior once again. Finally, the same picture is observed through MRE (Mean Relative Error) values, signifying small deviations between predicted and actual values and favouring MEKA over MLMAP.

A point that we feel is worth noting, however, is the following: The forecasting performance of both algorithms is considerably reduced in cases of a five-day prediction horizon, as expected. Nevertheless, the MLMAP algorithm exhibits a much preferable predictive ability when moving to the five-day horizon case, retaining a NRMSE slightly lower than unity and an improved MRE over the corresponding MEKA statistics. This indicates that the MLMAP may be more applicable for a longer-time prediction horizon, once the shock flag is replaced by the daily foreign exchange reserves used to defend the drachma rates.

Figures 4(a),(b) present the interest rates prediction results in both the daily and five-days horizon in graphical form for those architectures that performed best (4(a) MEKA, 4(b) MLMAP). The area on the left of the dashed line represents the training phase, while the one on the right the testing phase.

### 5.2. Artificial Shocks

Aiming at improving the performance of the algorithms, we have tried several additional runs introducing artificial shock periods in the training phase of the interest-rate series, thus providing the algorithm with more information concerning a typical reaction of the authorities to an attack against the domestic currency (Fig. 5). In other words, we have simulated the behaviour of the overnight rates during periods of pressure against the drachma, substituting the artificial rates that the simulation has produced for the actual rates during the specific, shock-free period of the training set. To avoid a favourable bias concerning the forecasting performance of the algorithms used, the artificial shocks do not resemble any particular crisis pattern as far as the overnight ($o/n$) rate fluctuations are concerned. Instead, they are designed to exhibit a typical crisis pattern, with the $o/n$ rate fluctuations being substantially pronounced and their pattern exhibiting two or more recurrent peaks as a result of the sustained pressure on the drachma. Selected results of these experiments are summarised in Table III. We observe that the introduction of an artificial shock

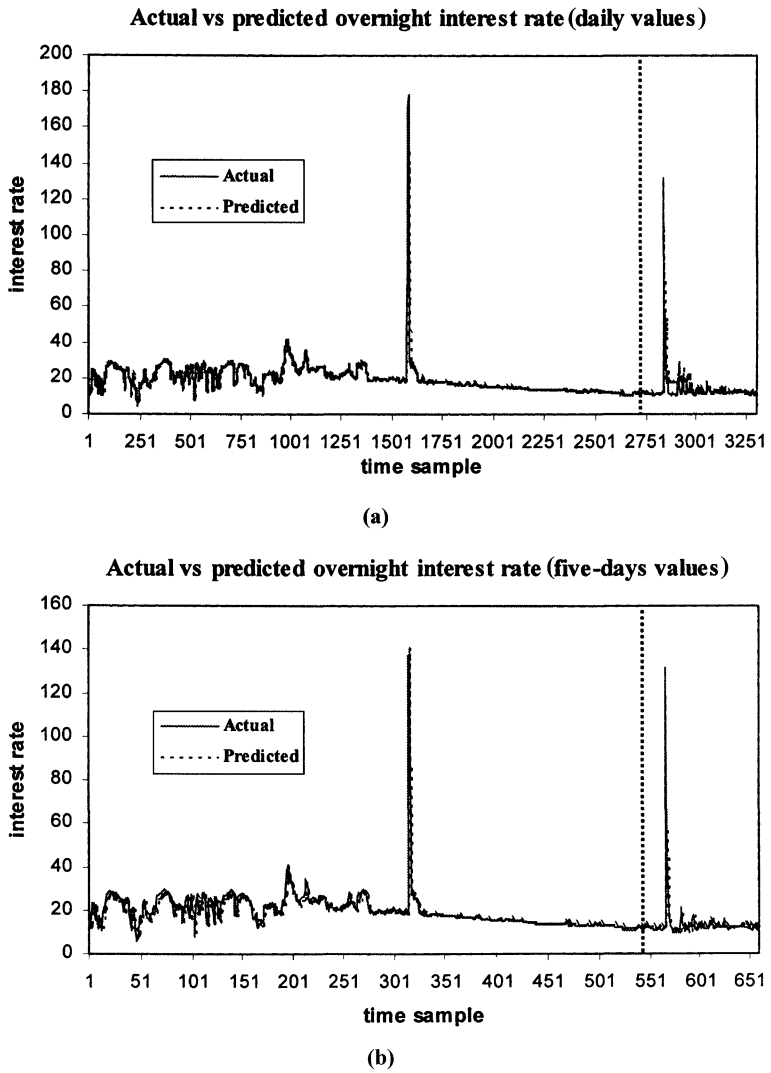



FIGURE 4 Actual *versus* predicted interest rate values produced by neural networks (a) MEKA, daily and (b) MLMAP, 5-days prediction horizon.

has improved the predictive ability of the MLMAP in terms of the NRMSE and, more so, in terms of the MRE, showing a remarkable improvement compared to the case in which artificial shocks were absent. It is interesting to point out that the forecasting performance of the algorithm does not decrease when the prediction horizon is increased from one to five days. The MEKA, on the contrary, does not appear to be suitable for such exercises, as both the NRMSE and the MRE deteriorate due to the introduction of the artificial shocks. The CC, finally, are somewhat lower in the artificial shocks case, reaching rates of slightly less than 70% for a one-day horizon, in the MLMAP case. A possible explanation for this inferior performance may be related to planting an artificial shock which assumes a more or less arbitrary pattern resembling none of those characterising the actual shock periods, thus, confusing the network rather than

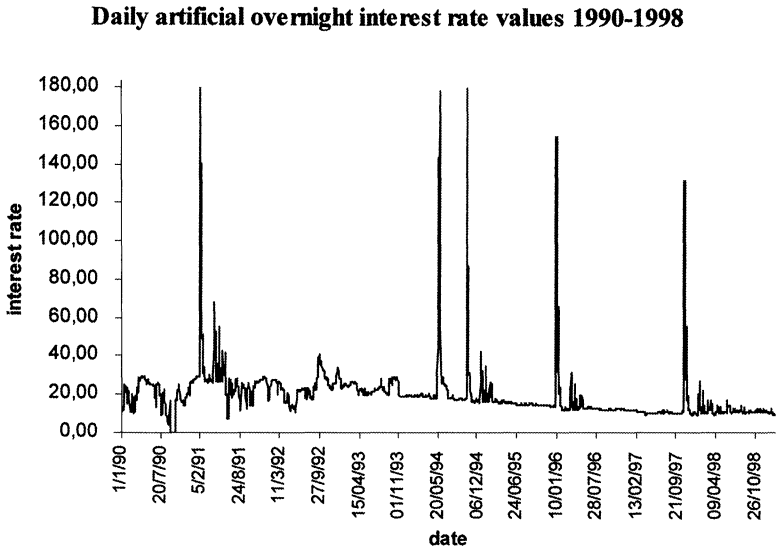



FIGURE 5 Daily values for the Greek overnight interest rate enriched with periods of artificial attacks against the drachma.

TABLE III Evaluation of artificial interest-rate prediction during the training and testing phases for a one and five-days prediction horizon

| | | Training phase | | | Testing phase | | |
|---|---|---|---|---|---|---|---|
| Horizon | Network architecture* | C.C. | NRMSE | MRE | C.C. | NRMSE | MRE |
| 1 | MEKA 8-15-1 | 0.8744 | 0.5013 | 0.0418 | 0.6577 | 1.1568 | 0.1236 |
| 5 | MEKA 7-5-1 | 0.8051 | 0.7280 | 0.1020 | 0.2024 | 1.5355 | 0.2555 |
| 1 | MLMAP 2-10-10-18-1 | 0.852 | 0.524 | 0.046 | 0.669 | 0.760 | 0.066 |
| 5 | MLMAP 2-10-10-18-1 | 0.882 | 0.780 | 0.105 | 0.273 | 0.989 | 0.128 |

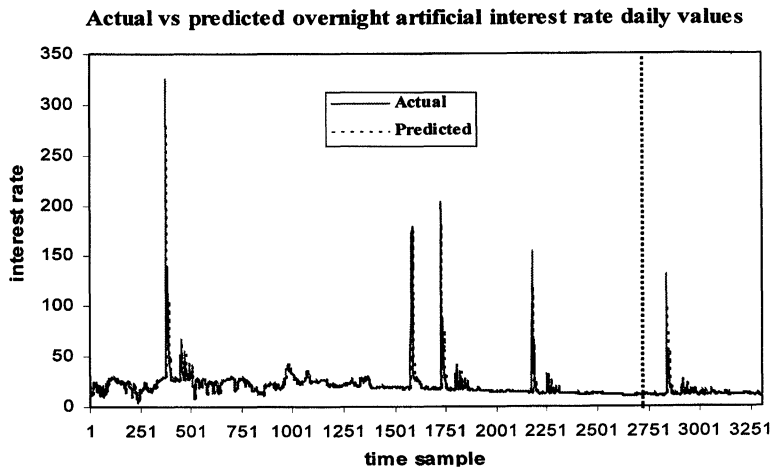
*MEKA $p$-$k$-$t$ refers to MEKA algorithm on a $p$-inputs, $k$-hidden nodes and $t$ outputs network MLMAP $p$-$k_1$-$k_2$-$k_3$-$t$ refers to MLMAP scheme with $p$-inputs, $k_1$ nodes in the first hidden layer, $k_2$ nodes in the second hidden layer, $k_3$ nodes in the third hidden layer and $t$ outputs.

improving its level of knowledge gained through training.

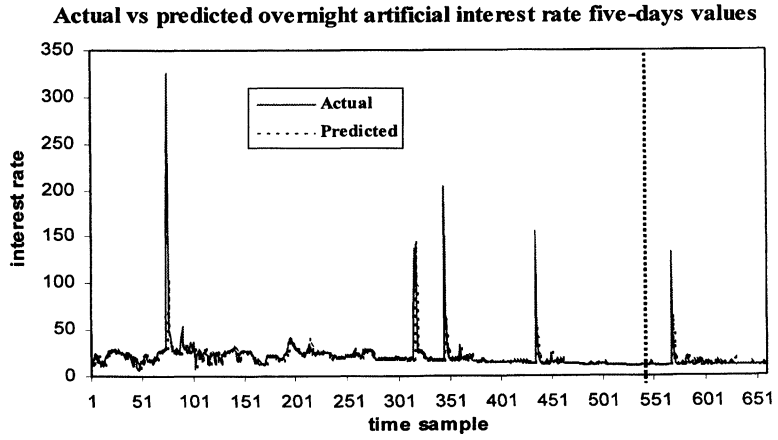
On the basis of the above and in terms of a general assessment, we are inclined to support the view that the MLMAP is considered more applicable for tackling forecasting exercises of this type. Indeed, Figures 6(a),(b) indicate the best results obtained using the MLMAP in terms of forecasting performance with the help of artificial interest rates, in both the daily and five-days horizon.

In terms of evaluating our results with reference to similar research on the topic, comparison can be made only with papers using time-series rather than model-driven techniques, since the success of the latter depends heavily on the performance of the function(s) used. Such functions concentrate on forecasting a shock based on the exchange-rate behaviour itself rather than the authorities reactions during an attack. This, in its turn, presupposes that the pattern of behaviour suggested by the function is expected to continue throughout the forecasting horizon, which is not always the case. Moreover, not intending to underestimate any such work, we need to point out that most of these contributions rely on the ERM participation of the currencies involved. This makes the prediction of an oncoming shock easy to determine, once the forecasted rates exceed the ERM band limits.

Turning to papers using data driven approaches, Steurer (1995) concentrates on forecasting the DEM/USD exchange rate using as input the daily spot rates. Despite the distinct differences between our approach and the one under consideration, one must point out that the accuracy rate of




(a)




(b)

FIGURE 6 Actual *versus* predicted interest rate values produced with the MLMAP algorithm using artificial series on a (a) daily and (b) 5-days prediction horizon.

fluctuation prediction ranges on the average around 50%, an accuracy rate typical of the so-called "naive prediction" method. As regards Andreou *et al.* (1998b), they apply the neural network methodology aiming at predicting a run against the drachma, using as input both the drachma rates against four major currencies, as well as the overnight rates. The results obtained on the basis of the MLP algorithm lead to a better forecasting performance compared to the present paper. One must point out, however, that the MLP uses a dual-nature shock flag, (0 or 1), lying on the output side, which facilitates the prediction performance, unlike the present paper, in which we use the three-parametric shock flag (0, 0.5, 1) as one of the inputs to predict the overnight interest rate fluctuations.

## 6. CONCLUSIONS

The results obtained in this paper lead to the conclusion that the neural network algorithms are indeed applicable in the context of a discrete dynamics environment, for predicting an oncoming attack against a certain currency. Despite the

lack of data concerning daily foreign-exchange expenditure by the Central Bank in defence of the domestic currency rates, the forecasting performance of the networks is quite satisfactory, based only on the daily interbank overnight interest rates which is the second policy instrument used by the authorities in cases of a crisis.

Two alternative computational intelligent systems based on MLP neural networks were employed to predict an oncoming shock relying primarily on the overnight interest rate fluctuations: A hybrid scheme trained with Kalman filtering and evolved by a genetic algorithm (MEKA) and a multiply activated architecture (MLMAP). Both schemes yielded encouraging results along a daily and five-day prediction horizon. Between the two algorithms employed in this paper, the MLMAP seems to be more suitable for such forecasting exercises compared to the MEKA. This conclusion has been reached on the basis of the improved learning ability displayed by the MLMAP, once the raw data is subjected to an environment of simulated attacks against the domestic currency, (the "artificial" crises expressed by the rise of the overnight rates). The MLMAP has also shown better performance in comparison to the MEKA in cases of longer-than-one-day forecasts.

The results above must be interpreted very carefully, since the forecasting performance of the networks can not be considered as predetermining the intentions of the authorities to resort to a devaluation in order to ease the pressure exercised on the domestic currency. Such reactions are the result of policy decisions which may be taken by the authorities either in an environment of speculative attacks on the currency, or just as a measure of improving export-price competitiveness, in the absence of a crisis altogether. The forecasting performance of the algorithms employed, therefore, simply shows that the system has been instructed to recognise certain signals which indicate that the currency is under considerable pressure which may or may not lead to a devaluation. In other words the predictive performance of the networks acts as a reliable hedging device rather than a means for an exchange-rate forecast. One should not overlook, however, useful indications reflected in the behaviour of the $o/n$ rates series, like the length of the time period during which the pressure is going on, as well as the persistent and recurrent pattern of abrupt overnight-interest rate increases. These may be directly related to the possibility of a devaluation, since they indicate that the measures taken may not be effective enough to mitigate or even curb the pressure on the currency.

Further research on this topic is expected to lead to more successful results, particularly with reference to the correlation coefficient, once data on the daily foreign-exchange expenditure during crises becomes available to replace the shock flag used as a proxy by the networks.

### *Acknowledgements*

The authors are indebted to Professor A. Bountis of the Mathematics Department, University of Patras, for valuable assistance during the preparation of the paper.

### *References*

Adamopoulos, A., Andreou, A., Georgopoulos, E., Ioannou, N. and Likothanassis, S. (1997) *Currency Forecasting Using Recurrently RBF Networks Optimized by Genetic Algorithms*. Computational Finance 1997 (CF'97), London Business School.

Andreou, A., Georgopoulos, E., Likothanassis, S. and Polidoropoulos, P. (1997) Is the Greek Foreign Exchange-rate Market Predictable? A Comparative Study Using Chaotic Dynamics and Neural Networks. *Proceedings of the Fourth International Conference on Forecasting Financial Markets*, Banque Nationale de Paris and Imperial College, London.

Andreou, A., Georgopoulos, E., Zombanakis, G. and Likothanassis, S. (1998a) Testing Currency Predictability Using an Evolutionary Neural Network Model. *Proceedings of the Fifth International Conference on Forecasting Financial Markets*, Banque Nationale de Paris and Imperial College, London.

Andreou, A., Zombanakis, G., Georgopoulos, E. and Likothanassis, S. (1998b) Modeling and Forecasting Exchange-rate Shocks. *Proceedings of 60eme International Conference on New Financial Instruments and Market Localisation*, Applied Econometrics Association (AEA), Paris.

Andreou, A. S., Karytinos, A. and Pavlides, G. (2000) Non-linear Time-series Analysis of the Greek Exchange-rate Market. *International Journal of Bifurcation and Chaos*, **10**(7) (forthcoming).

Bank of Greece (1991) *Report of the Governor*, Athens.

Bank of Greece (1994) *Report of the Governor*, Athens.

Bank of Greece (1997) *Report of the Governor*, Athens.

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

Berg, A. and Pattillo, C. (1999) Are Currency crises Predictable? A Test. *IMF Staff Papers*, **46**(2), 107–138.

Brissimis, S. N. and Leventakis, J. A. (1989) The Effectiveness of Devaluation: A General Equilibrium Assessment with Reference to Greece. *Journal of Policy Modeling*, **11**(2), 247–271.

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

Frankel, J. A. (1993) *On Exchange Rates*. MIT Press.

Haykin, S. (1994) *Neural Networks, A Comprehensive Foundation*. McMillan College Publishing Company, New York.

Karadeloglou, P. (1990) On the Existence of an Inverse *j* Curve. *Greek Economic Review*, **12**, 285–305.

Karadeloglou, P., Papazoglou, C. and Zombanakis, G. A. (1998) Is the Exchange Rate an Effective Anti-inflationary Policy Instrument? *Economia*, **2**(1), 47–72.

Karytinos, A., Andreou, A. S. and Pavlides, G. (1999) Long-term Dependence in Exchange Rates. *Journal of Discrete Dynamics in Nature and Society*, **3**(1).

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

Koedijk, K. and Kool, C. (1994) Tail Estimates and the EMS Target Zone. *Review of International Economics*, **2**(2), 1153–1165.

Leventakis, J. (1987) Exchange-Rate Models: Do they Work? *Weltwirtschaftliches Archiv*, **123**, 363–376.

Levich, R. (1989) Forward Rates as the Optimal Future Spot Rate Forecast. *Exchange-Rate Forecasting* (Dunis, C. and Feeny, M. Eds.). Probus, Chicago.

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.

Majuca, R. (1992) Balance-of-payments Crises: Timing and Collapse of the Philippine Peso. *Philippine Review of Economics and Business*, **29**(2), 231–250.

Marsh, I. W. and Power, D. M. (1996) A Note on the Performance of Foreign Exchange Forecasters in a Portfolio Framework. *Journal of Banking and 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. *Neural Networks in the Capital Markets* (Refenes, A. P. Ed.). Wiley, U.K.

Michalewicz, Z. (1996) *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag.

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.

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

Rose, A. and Svensson, L. (1995) Expected and Predicted Realignments: The FF/DM Exchange Rate During the EMS, 1979–1993. *Scandinavian Journal of Economics*, **97**(2), 173–200.

Shah, S., Palmieri, F. and Datum, M. (1992) Optimal Filtering Algorithms for Fast Learning in Feed-forward Neural Networks. *Neural Networks*, **5**, 779–787.

Siklos, P. and Tarajos, R. (1996) Fundamentals and Devaluation Expectations in Target Zones: Some New Evidence from the ERM. *Open Economics Review*, **7**(1), 35–59.

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

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

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

Verrier, M. (1989) Selection and Application of Currency Forecasts. Exchange-rate Forecasting (Dunis, C. and Feeny, M. Eds.). Probus, Chicago, 303–343.

Zombanakis, G. A. (1998) Is the Greek Exporters Price Policy Asymmetric? *Greek Economic Review*, **19**(1).