Improved neural network forecasting models for short term foreign exchange rates using volatility indices

Hassan Dargahi

Faculty of Economic and Political Sciences
Shahid Beheshti University
Tehran, Iran

h-dargahi@sbu.ac.ir
h-dargahi@sharif.ac.ir

Tel: (9821) 66 16 58 61
Fax: (9821) 66 02 27 59

For presentation at the
2008 Business & Economics Society International Conference
July 15-19, 2008
Lugano Switzerland, Hotel De La Paix

SUBJECT CATEGORY TITLE AND NUMBER:
FINANCE & INTERNATIONAL FINANCE (F4)
QUANTITATIVE METHODS (QM6)

Improved neural network forecasting models for short term foreign exchange rates using volatility indices

Abstract

The emphasis of this paper is the role of volatility indices on improvement Artificial Neural Networks (ANNs) forecasting models for the daily USD/EUR and USD/GBP exchange rates Two volatility indices are used. The realized volatility which is based on intra-daily data and the GARCH volatility. They are applied into the model in two ways. Firstly, the lagged volatility index is added to the model. Secondly, some levels for the volatility are defined and the time series are partitioned according to the level of volatility on time t, and then different models of exchange rate forecasting on time t+1 are built for each level of volatility.

The forecasting results demonstrate that the models with low and middle volatility are not preferred to the model without volatility index. But in case of high volatility, the level models improve forecasting power. This means that high volatility is new information for foreign exchange market

.

***Keywords**: exchange rate, forecasting, neural network, volatility*

1 Application of ANNs for Exchange Rates Forecasting

Forecasting exchange rates is an important financial problem that is receiving increasing attention especially because of its difficulty and practical applications. Over the past few years, Artificial Neural Networks (ANNs) have been widely advocated as a new alternative modeling technology to more traditional econometric and statistical approaches, claiming increasing success in the fields of economic and financial forecasting. This has resulted in many publications comparing neural networks and traditional forecasting approaches.

Conventional time series models rely on global approximation, and are well suited to problems with stationary dynamics. In the analysis of real world systems, however, two of the key problems are non stationarity (often in the form of switching between regimes) and overfitting (which is particularly serious for noisy processes) (Weigend et al. [1]). Non stationarity implies that the statistical properties of the data generator vary through time. This leads to gradual changes in the dependency between the input and output variables. Noise, on the other hand, refers to the unavailability of complete information from the past behavior of the time series to fully capture the dependency between the future and the past. Noise can be the source of overfitting, which implies that the performance of the forecasting model will be poor when applied to new data (Cao [2]; Milidiu & Renteria [3]).

Artificial neural networks (ANNs) have been widely used as a promising alternative approach for a forecasting task because of several distinguished features. Research efforts on ANNs for forecasting exchange rates are considerable. An ANN is a system loosely modeled on the human brain, which detect the underlying functional relationships within a set of data and perform tasks such as pattern recognition, classification, evaluation, modeling, prediction and control. ANNs are particularly well suited to finding accurate solutions in an environment characterized by complex, noisy, irrelevant or partial information. Several distinguishing features of ANNs make them valuable and attractive in forecasting. First, as opposed to the traditional model-based methods, ANNs are data-driven self-adaptive methods in that there are few *a priori* assumptions about the models for problems under study. Second, ANNs can generalize. Third, ANNs are universal functional approximators. Finally, ANNs are nonlinear.

The idea of using ANNs for forecasting exchange rates is not new. Weigend *et al* [1] find that neural networks are better than random walk models in predicting the DEM/USD exchange rate. Kuan and Liu [4] use both feed-forward and recurrent neural networks to forecast GBP, CAD, DEM, JPY, CHF against USD. Hann and Steurer [5] make comparisons between the neural network and linear model in USD/DEM forecasting. Refense *et al.* [6] apply a multi-layer perception network to predict the exchange rate between USD/DEM, and to study the convergence issue related to network architecture. Refense [7] develops a constructive learning algorithm to find the best neural network configuration in forecasting DEM/USD. Podding [8] studies the problem of predicting the trend of the USD/DEM, and compares results to those obtained through regression analysis. Pi [9] proposes a test for dependence among exchange rates. Shin [10] applies an ANN model and moving average trading rules to investigate return predictability of exchange rates. Zhang and Hutchinson [11] report the experience of forecasting the tick-by-tick CHF/USD. Wu [12] compares neural networks with ARIMA models in forecasting Taiwan/USD exchange rates. Dunis [13] investigated the application of NNR to intraday foreign exchange forecasting and his results were evaluated by means of a trading strategy. Episcopos and Davis [14] investigate the problem of predicting daily returns based on five Canadian exchange rates using ANNs and a heteroskedastic model, EGARCH. Tenti [15] applied ANNs to predict the USD/DEM exchange rate, devising a trading strategy to assess his results, while Franses and Homelen [16] use ANN models to predict four daily exchange rate returns relative to the Dutch guilder using directional accuracy to assess out-of-sample forecasting accuracy. Other examples using ANN in exchange rates application include Zhang [17] and Yao *et al* [18]. They provide a brief description of neural networks, their advantages over traditional forecasting models, and their applications for business forecasting.

Although global approximation methods can be applied to model and forecast time series having the aforementioned characteristics, it is reasonable to expect that the forecasting accuracy can be improved if regions of the input space exhibiting similar dynamics are identified and subsequently a local model is constructed for each of them. A number of researchers have proposed alternative methodologies to perform this task electively (Cao [2]; Milidiu & Renteria [3], 1999; Pavlidis et al. [19] ; Pavlidis et al. [20]; Principe et al., 1998; Sfetsos & Siriopoulos [22]; Weigend et al.[23]). In principal, these methodologies are formed by the combination of two distinct approaches; an algorithm for the partitioning of the input space and a function approximation model. Evidently the partitioning of the input space is critical for the successful application of these methodologies. Dunis and Huang [24] examine the use of GARCH models, Neural Network Regression (NNR), Recurrent Neural Network (RNN) regression and model combinations for forecasting and trading currency volatility, with an application to the GBP/USD and USD/JPY exchange rates. Both the results of the NNR/RNN models and the model combination results are benchmarked against the simpler GARCH alternative. Huang et al [25] compare the predication performance of neural networks with the different frequencies of input data, namely daily data, weekly data, monthly data. In the 1 day and 1 week ahead prediction of foreign exchange rates forecasting, the neural networks with the weekly input data performs better than the random walk models. In the 1 month ahead prediction of foreign exchange rates forecasting, only the special neural networks with weekly input data perform better than the random walk models. Because the weekly data contain the appropriate fluctuation information of foreign exchange rates, it can balance the noise of daily data and losing information of monthly data.

In this paper I investigate the improvement of neural network forecasting models for short term (daily) foreign exchange rates using volatility indices. The remaining paper is organized as follows: in the next section ANNs by the use of volatility indices are developed. In section 3 empirical results regarding the spot exchange rate of the USD/EUR and USD/GBP are presented. The paper ends with a short discussion of the results and concluding remarks.

2 ANNs Development by the Use of Volatility Indices

What makes this study different from the other empirical works is the use of exchange rate volatility indices to improve the accuracy of forecasting in the technical models of ANNs.

2.1 Input Selection

The inputs used for models building are USD/EUR and USD/GBP exchange rates. The large sample consists of 1363 observation from 2001 to 2006. In this data set, the inputs are available during 15 minutes and for the 24 hours of the day. The 15 minutes input during a day is used for measuring the daily volatilities. The exchange rate of the hour 24 according to the international time is considered as the representative of that day’s exchange rate. The Matlab software is also used for the programming and design of the neural networks.

2.1.1 Stationary Test

According to the ADF test, the daily time series of USD/EUR is not stationary in the 10% critical level. Transforming the data to the growth rate time series related to the average of past 5 days (the average of last working week) will make the time series stationary. The reason for this transform is that the modeling based on the daily exchange rate input which is highly volatile, will not result the significant estimations. The imposed changes on the daily exchange rate are done based on following relations:

s_i : daily exchange rate (1)

$$MA_i = \sum_{j=1}^{5} s_{i-j+1} \qquad i = 5,6,...1363$$

$$r_i = \log s_{i+1} - \log MA_i \quad i = 1,2,3,...,1363 \tag{2}$$

2.2.2 Volatility Indices

What is observed about variables such as stock return or exchange rate is that changes are a lot in some periods and few in others. Besides, when the changes are high or low, this condition will continue for a while. Thus, volatility is time varying and has a clustering characteristic. This means that low (high) changes will be followed with low (high) changes for a considerable period of time. In other words volatility is an autoregressive and stationary.

2.2.2.1 Variance Volatility Index By using of data with high frequency, a volatility index named the realized volatility can be computed. There are 96 data every day with 15 minutes intervals. The sum of squared of the 15 minutes growth rate data is the realized volatility. This index is hereafter called the Variance Volatility.

2.2.2.2 GARCH Volatility Index There is also another way for measuring the volatility. The process of generating the return time series (r_i) is considered as the GARCH (1, 1) and the conditional variance time aeries of the residual terms is derived as the GARCH Volatility.

2.2.3 Lags Structure

For choosing the lags of a time series model, the autocorrelation coefficient criterion (ACC), unlike the Akaike or Bayesian criteria, has no presumption about the forecasting variable. Huang et al [26] have reported that the forecasting accuracy of the daily exchange rate based on autocorrelation to determine input variables of neural networks is better than the two other ways. The same way also used in this study for choosing the lags structure. The logic of this approach is as following:

- The absolute value of the autocorrelation coefficient must be the most between the lagged and spot exchange rate..
- The sum of absolute value of autocorrelation coefficients between the lag and other chosen lags must be the least.

By assuming the influence of the lags till to 10th one, the lag structure has been determined as {1, 5, 7, 9}.

2.2.4 Definition of Volatility Thresholds

The focus of this study is to separate the inputs based on the volatility level, and to develop different forecasting ANNs for each grouped input data. For partitioning the inputs, the volatility thresholds are chosen in a way that the numbers of observations are the same in each model. For example, for separating the inputs to 3 levels, the first threshold is a number that one third of the observations have a lower volatility than it, and the second threshold is a number that one third of the observations have volatility between the first and second thresholds.

2.3 Training, Test and Validation Sets

It is important to note that in data splitting, the issue is not about what proportion of data should be allocated in each sample. But, rather, it is about sufficient data points in each sample to ensure adequate learning, validation, and testing. Granger [27] suggests that for nonlinear modeling at least 20% of the data should be held back for an out-of-sample evaluation. Hoptroff [28] recommends that at least 10 data points should be in the test sample while Ashley [29] suggests that a much larger out-of-sample size is necessary in order to achieve statistically significant improvement for forecasting problems. In this study the training, test and validation sets were partitioned approximately 70%, 20%, and 10% of the input / output pairs of each level, respectively. Training set includes the oldest inputs and validation set contains the recent ones. The union of the validation sets of all levels makes the validation set of the base model (with no volatility index). The test and validation sets of the base model are made in the same way.

2.4 ANN Architecture Selection and Training Algorithm

Numerous neural network models have been proposed, but multilayered Feed-forward Neural Networks (FNNs) are the most common. In FNNs neurons are arranged in layers and there are connections between neurons in one layer to the neurons of the following layer. In this research I employ feed-forward multilayer perceptron ANN models with two hidden layers. The networks have 4 or 5 neurons in the first layer and one in the last, depending on the number of input variables. The number of neurons in the two middle layers differs. As the output of the model is the exchange rate return and quantified between -1 and 1, then the activation function is a tangent sigmoid hyperbolic. This function is also used for the neurons of the hidden layers.

The most widely used training method for ANN models is the error back propagation

(BP) algorithm, which is a recursive gradient descent method, where the network weights θ are chosen to minimize a loss function, typically the sum of squared errors. The standard back propagation algorithm is often too slow for practical problems. Therefore, a notable faster variation of the BP algorithm, namely the Levenberg-Marquardt (LM) algorithm, was used. The main difference between them is that the LM algorithm uses an approximation of the Hessian matrix.

The widest network which is known in other studies for exchange rate forecasting is a network with 6 neurons in the first layer and 4 in the second. Therefore searches for optimal structures are limited to networks with 1 to 6 neurons in the first layer and 1 to 4 in the second. For selecting a network, 24 different structures are trained each time and the network with the less out-of-sample error is selected among them. This is repeated 10 times with initial random weights and then the network with the most determination coefficient or less in-sample error is selected.

3 Presentation of Empirical Results

3.1 Base Model

The first trained network is the one which its inputs are 4 lagged daily exchange rates, with no entered volatility index. This model is called the base model and its result is shown in Table (1). The mentioned network has 4 neurons in the first layer and one in the second. In this case, the Mean Squared Error normalized with variance (NMSE) is 0.55. This means that the coefficient of determination is 45 percent. Following equations explain variables of the above table:

$$NMSE = \frac{\frac{1}{N}\sum(r_t - \hat{r}_t)^2}{\mathrm{var}(r_t)} \; ; \; RMSE = \sqrt{\frac{1}{N}\sum(r_t - \hat{r}_t)^2} \; ; \; MAE = \frac{1}{N}\sum\left|r_t - \hat{r}_t\right| \; ; \; MAY = \frac{1}{N}\sum|r_t|$$

Table 1: Results for the base model

Model	Structure	NMSE (R^2)	DW	LWG	In-sample		Out-of-sample	
					MAE (MAY)	RMSE	MAE (MAY)	RMSE
Base	4×1	0.55 (0.45)	1.96	0.1000	0.0051 (0.0072)	0.0066 (0.0089)	0.0039 (0.0057)	0.0052 (0.0073)

DW: Durbin Watson statistic; LWG: Li-White-Granger

3.2 Model with lagged variance volatility

In this model the first lag of the variance volatility is added to the 4 lagged exchange rate variables as an input. The results are reported in Table (2). The coefficient of determination has increased only 2 percent in comparison with the base model and the out-of-sample RMSE has decreased from 0.0052 to 0.0050.

Table 2: Results for the model with a lagged variance volatility

Model	Structure	NMSE (R^2)	DW	LWG	In-sample		Out-of-sample	
					MAE (MAY)	RMSE	MAE (MAY)	RMSE
Variance volatility as input	6×4	0.53 (0.47)	1.96	0.1000	0.0050 (0.0072)	0.0066 (0.0089)	0.0038 (0.0057)	0.0050 (0.0073)

3.3 Models of variance volatility levels: (Separating the inputs according to the variance volatility level)

Now the inputs are separated into three categories based on the first lag of the variance volatility and models are estimated for each group of inputs separately, as follows:
First level model: low volatility
Second level model: middle volatility
Third level model: high volatility

The results show (Table 3) that by increasing the volatility level, the coefficient of determination increases from 26% in the first model to 46% and 58% in the second and third models, respectively. For considering the effect of number of input categories on the network performance, this time the inputs are separated to 5 levels. In this case the coefficient of determination is increased from 11% to 66% in an ascending trend. The coefficients of determination of the forth and fifth level model are more than the base model.

Table 3: Results for the models with separated inputs based on three variance volatility levels

Model	Structure	NMSE (R^2)	DW	LWG	In-sample		Out-of-sample	
					MAE (MAY)	RMSE	MAE (MAY)	RMSE
Leve l One	4×1	0.74 (0.26)	2.00	0.1000	0.0055 (0.0066)	0.0070 (0.0081)	0.0044 (0.0052)	0.0057 (0.0066)
Level Two	5×3	0.54 (0.46)	1.89	0.1000	0.0051 (0.0071)	0.0065 (0.0089)	0.0031 (0.0047)	0.0043 (0.0060)
Level Three	3×2	0.42 (0.58)	2.02	0.1000	0.0049 (0.0081)	0.0064 (0.0198)	0.0036 (0.0065)	0.0047 (0.0081)

3.4 Comparison between models of variance volatility levels and base model

In this section the forecasting error of models with the same validation set are compared. It means that the validation set of each model of variance volatility level is given to the base model and its forecasting error is compared with the out-of-sample error of the former models. The results demonstrate (Table 4) that in all cases the out-of-sample error of models for variance volatility levels has decreased. For evaluating the significance of the prediction improvement of the volatility level models against the base model, the Diebold-Marino [30] test is used. The results show that the first level model has not been preferred to the base model but the second and third level models, in the 76 and 72 percent level, have a significant improvement in comparison with the base model (Table 5). The above test is repeated for the case of five variance volatility level models. According to the Diebold-Marino test, the forth and fifth model in the 78 and 77 percent level have a significant improvement in comparison with the base model.

Table 4: Out-of-sample error of the three variance volatility level models and base model with the same validation set

Data	Out-of-sample error of the base model		Out-of-sample error of the level model	
	MAE (MAY)	RMSE	MAE (MAY)	RMSE
Level One	0.0044 (0.0052)	0.0058 (0.0066)	0.0044 (0.0052)	0.0057 (0.0066)
Level Two	0.0033 (0.0047)	0.0045 (0.0060)	0.0031 (0.0047)	0.0043 (0.0060)
Level Three	0.0038 (0.0065)	0.0049 (0.0081)	0.0036 (0.0065)	0.0047 (0.0081)

Table 5: Test for prediction performance of the three variance volatility level models

Model	Diebold-Marino	Level of significance (%)
Level One	0	50
Level Two	-0.71	76
Level Three	-0.59	72

3.5 Model with lagged GARCH volatility

The above ANN modeling is continued with the second definition of volatility that is the Garch volatility index. The model in which the first lag of the Garch volatility, as an input, is added to the four lagged variables of the exchange rate. In this case the coefficient of determination is 48% which has no obvious increase in comparison with the 45% of the base model. (Table 6).

Table 6: Results for the model with a lagged Garch volatility

Model	Structure	NMSE (R^2)	DW	LWG	In-sample		Out-of-sample	
					MAE (MAY)	RMSE	MAE (MAY)	RMSE
GARCH volatility as input	7×6	0.52 (0.48)	1.98	0.1000	0.0050 (0.0072)	0.0065 (0.0089)	0.0039 (0.0057)	0.0052 (0.0073)

3.6 Models of GARCH volatility levels: (Separating the inputs according to the Garch volatility level)

First, the inputs are divided into three groups of low, middle and high volatility according to the first lag of the GARCH volatility. Then, models are estimated for each group of inputs separately. The results show that, as before, by increasing the volatility level, the coefficients of determination of ANNs increase. For considering the effect of number of input categories on the network performance, this time the inputs are separated to 5 volatility levels. In this case the coefficient of determination is increased from 31% to 58% in an ascending trend. The coefficients of determination of the forth and fifth level model are more than the base model.

3.7 Comparison between models of GARCH volatility levels and base model

For this comparison, the forecasting errors of models with the same validation set are considered. It means that the validation set of each model of GARCH volatility level is given to the base model and its forecasting error is compared with the out-of-sample error of the former models. The results for models based on three volatility levels demonstrate that in all cases the out-of-sample error of models for GARCH volatility levels has decreased. According to the Diebold-Marino test, the high volatility model (level three) in a 81% level, has a significant improvement in comparison with the base model. The above test is repeated for the case of five GARCH volatility level models. Results of the Diebold-Marino test show that the forth and fifth model in the 99 and 81 percent level have a significant improvement in comparison with the base model.

4 Concluding Remarks

In this paper the modeling of Neural Network is used for forecasting of the daily USD/EUR and USD/GBP exchange rates. What makes this study different from the other empirical works is the use of exchange rate volatility indices to improve the accuracy of forecasting in the technical models of ANNs. Two volatility indices are employed. The realized (or variance) volatility which is based on intra-daily data and the GARCH volatility. They are applied into the models in two ways. Firstly, the lagged volatility index is added to the model. Secondly, some levels for the volatility are defined and the inputs are separated according to the level of volatility, and then different models of exchange rate forecasting are built for each level of volatility. The results of the study can be summarized as follows:

- The addition of the first lag of the volatility index to the model's inputs makes no significant improvement in prediction of the model in comparison with the base model (with no volatility index). This shows that the volatility of each day has no effect on expectations of tomorrow's exchange rate and for a market analyst who uses spot exchange rate, today's volatility is not known as new information for improvement of prediction.
- The results of the forecasting models by considering the leveling of the volatility index show that the higher volatility level models improve the forecasting accuracy in comparison with the base model but there is no such improvement in the middle and low levels. This means that in the exchange market the higher volatility is known as a shock and new information so that it is not exist in information set of players for expectations forming. Therefore, adding it to the model improves the exchange rate prediction. But the lower volatility is expected by the players so adding it to the model dose not give more information than the lagged exchange rates.
- Separation of input data into more categories based on different volatility levels result the better performance for the forecasting ANNs models with high volatility levels.
- The results are not sensitive to the definition of the volatility index. Because by considering the use of two volatility indices (variance and GARCH volatilities), forecasting accuracy of the models are too much alike.
- For studying the sensitivity of the models performance to the selected exchange rate, development of all models is repeated in a same way for the USD/GBP. In this case also the results are similar. In other words, the short term expectations of the players of different foreign exchange markets are the same according to the information related to the volatility level.

References

[1] Weigend, A.S. & Huberman, B.A. & Rumelhart, D.E., Predicting sunspots and exchange rates with connectionist networks, in *Nonlinear Modeling and Forecasting*, eds. M. Casdagli and S. Eubank, pp. 395–432, Addison-Wesley, Redwood City, CA, 1992.

[2] Cao, L., Support vector machines experts for time series forecasting. *Neurocomputing*, 51, pp. 321-329, 2003.

[3] Milidiu, R. L., Machado, R. J. & Renteria, R. P., Time-series forecasting through wavelets transformation and a mixture of expert models. *Neurocomputing*, 28, pp. 145-156, 1999.

[4] Kuan, C.M. & Liu, T., Forecasting exchange rates using feedforward and recurrent neural networks. *Journal of Applied Econometrics*, 10, pp. 347–364, 1995.

[5] Hann, T.H. & Steurer, E., Much ado about nothing? Exchange rate forecasting: Neural networks versus linear models using monthly and weekly data. *Neurocomputing*, 10, pp. 323–339, 1996.

[6] Refenes, A.N., Azema-Barac, M., Chen, L. & Karoussos, S.A., Currency exchange rate prediction and neural network design strategies. *Neural Computing and Application*, 1, pp. 46–58, 1993.

[7] Refenes, A. N., Constructive learning and its application to currency exchange rate forecasting, in *Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real-World Performance*, eds. R. R. Trippi and E. Turban, Probus Publishing Company, Chicago, pp. 777–805, 1993.

[8] Podding, A., Short term forecasting of the USD/DEM exchange rate, in *Proceedings of First International Workshop on Neural Networks in Capital Market*, London, 1993.

[9] Pi, H., Dependence analysis and neural network modeling of currency exchange rates, *Proceedings of First International Workshop on Neural Networks in Capital Market*, London, 1993.

[10] Shin, S. H., *Essays on asset returns predictability*, PhD Dissertation, Massachusetts Institute of Technology, (0753), 1993.

[11] Zhang, X. & Hutchinson, J., Simple architectures on fast machines: Practical issues in nonlinear time series prediction, in *Time Series Prediction: Forecasting the Future and Understanding the Past*, eds. A. S. Weigend and N. A. Gershenfeld (Addison-Wesley, Reading, MA), pp. 219–241, 1994.

[12] Wu, B., Model-free forecasting for nonlinear time series (with application to exchange rates), *Computational Statistics and Data Analysis* 19, pp. 433–459, 1995.

[13] Dunis, C.L. & Williams, M., Modeling and Trading EUR/USD Exchange Rate: Do Neural Models Perform Better?. *Trading and Regulation*, 2002.

[14] Davis, J. T. & Episcopos, A. & Wettimuny, S., Predicting direction shifts on Canadian-US exchange rates with artificial neural networks. *International Journal of Intelligent Systems in Accounting, Finance and Management,* 10, pp. 83–96, 2001.

[15] Tenti, P., Forecasting foreign exchange rates using recurrent neural networks. *Applied Artificial Intelligence*, 10, pp. 567–581, 1996.

[16] Franses, P.H. & Homelen, P. V., On forecasting exchange rates using neural networks. *Applied Financial Economics,* 8, pp. 589-596, 1998.
[17] Zhang, G. P., An investigation of neural networks for linear time-series forecasting. *Computers & Operations Research,* 28, pp 1183-1202, 2001.
[18] Yao, J.T. & Tan, C.L., A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing*, 34, pp. 79–98, 2000.
[19] Pavlidis, N. G., Tasoulis, D. K. & Vrahatis, M. N., Financial forecasting through unsupervised Clustering and evolutionary trained neural networks. *Proceedings of the Congress on Evolutionary Computation*, pp. 2314-2321, 2003.
[20] Pavlidis, N. G., Tasoulis, D. K. & Vrahatis, M. N., Time series forecasting methodology for multiple step ahead prediction. *The IASTED International Conference on Computational Intelligence*, pp. 456-461, 2005.
[21] Principe, J. C., Wang, L. & Motter, M. A., Local dynamic modeling with self-organizing maps and applications to nonlinear system identification and control. *Proceedings of the IEEE*, 86(11), pp. 2240-2258, 1998.
[22] Sfetsos, A. & Siriopoulos, C., Time series forecasting with a hybrid clustering scheme and pattern recognition. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, 34(3), pp. 399-405, 2004.
[23] Weigend, A. S., Mangeas, M. & Srivastava, A. N., Nonlinear gated experts for time series: discovering regimes and avoiding overfitting. *International Journal of Neural Systems*, 6, pp. 373-399, 1995.
[24] Dunis, C. & Huang, X., Forecasting and Trading Currency Volatility: An Application of Recurrent Neural Regression and Model Combination. *The Journal of Forecasting,* 2001.
[25] Huang, W., Yu, L., Wang, S., Bao, Y., & Wang, L., Comparisons of the Different Frequencies of Input Data for Neural Networks in Foreign Exchange Rates Forecasting. in V.N. Alexandrov et al. (Eds.): ICCS, Part IV, LNCS 3994, pp. 517 – 524, 2006.
[26] Huang, W., Forecasting foreign exchange rates with artificial neural networks: a review. *International journal of Information Technology & Decision Making*, Vol. 3, No. 1, pp. 145–165, 2004.
[27] Granger, C. W. J., Strategies for modeling nonlinear time-series. *The Economic Record,* 69 (206), pp. 233-238, 1993.
[28] Hoptroff, R. G., The principles and practice of time series forecasting and business modeling using neural networks. *Neural Computing and Applications,* 1, pp. 59-66, 1993.
[29] Ashley, R, Statistically significant forecasting improvements: How much out-of-sample data is likely necessary?. *International Journal of Forecasting*, 2003.
[30] Diebold, F., Mariano, R., Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3), pp. 253– 263, 1995.