

Some Statistical Models vs. Models Based on SC Applied to Daily Exchange Rates*

Dusan Marcek

Research Institute of the IT4Innovations Centre of Excellence, The Silesian University in Opava, Czech Republic
dusan.marcek@fpf.slu.cz

Abstract - We developed forecasting models based on the advanced statistical (stochastic) and soft computing (computational intelligence) techniques for predicting high frequency data sets. Firstly, we used the standard statistical tools such as the autocorrelation/partial autocorrelation function, clustering, etc. to identify the chunks of information that are deemed essential for knowledge representation. Afterwards, (1) we proposed statistical methods to identify the relationship between the information granules; (2) based on the platform of granular or soft computing (G or SC) we developed and formal expressed the underlying mechanisms (models) that generate the observed data and, in turn, to forecast future values of the investigated process in managerial decision-making. The proposed intelligent approach is applied to the time series of USD/EUR exchange rates. We found that it is possible to achieve significant risk reduction in managerial decision-making by applying intelligent forecasting models based on the latest information technologies. We show that statistical GARCH-class models can identify the presence of the leverage effect and to react to the good and bad news. In a comparative study is shown, that both presented modeling approaches are able to model and predict high frequency data with reasonable accuracy, but the neural network approach is more effective and accurate.

Index Terms – ARIMA/ARCH/GARCH models, forecast accuracy, soft computing, RBF neural networks.

I. Introduction

In economics and in particular in the field of financial markets, forecasting is very important because forecasting is an essential instrument to operate day by day in the economic environment. Over the past ten years academics of computer science have developed new soft techniques based on latest information technologies such as soft, neural and granular computing to help predict future values of high frequency financial data. At the same time, the field of financial econometrics has undergone various new developments, especially in finance models, stochastic volatility, and software availability. Recently, most developed statistical (econometric) models assume a nonlinear relationship among variables. As example are the exponential and power GARCH models and autoregressive models. These are model-driven approaches based on a specific type relation among the variables. Neural networks and other soft computing techniques, on the other hand, are data driven models and nonparametric models. Unlike in classical statistical inference, the parameters are not

predefined and their number depends on the training data used. Parameters that define the capacity of model are data-driven in such a way as to match the model capacity to the data complexity. In this paper, two novel forecasting models are proposed for EUR/USD exchange rate prediction. The first based on latest statistical methods makes use of ARIMA/GARCH-class models, and another is the neural network based on the radial basic activation function that makes uses both supervised learning methods and unsupervised learning methods. Then, we discuss certain management aspects of proposed forecasting models such as capabilities and interests of the people who will make and use the forecast, in their decision processes.

The paper is organized as follows. In Section 2 we briefly describe basic ARCH-GARCH models and characterize the neural approach. In Section 3 we present the data, conduct some preliminary analysis of the time series and demonstrate the forecasting abilities of ARCH-GARCH modes of an application. In Section 4 we introduce the architectures of RBF (Radial Basic Function) networks. In Section 5 we put an empirical comparison. Section 6 briefly concludes.

II. Theoretical Background

Traditional statistical/econometric models assume a constant one-period forecast variance. But, the financial time series features various forms of nonlinear dynamics, the crucial one being the strong dependence of the instantaneous variability of the series on its own past [1]. To predict the financial time series data a regression model is used with disturbances following an ARCH type process.

A. ARIMA Time Series Models

Time series models have been initially introduced either for descriptive purposes like prediction or for dynamic control. In this paper we will use linear time series models so-called ARIMA which are very easy implement well-established methods for time series prediction. They combine autoregressive (AR), and moving average (MA) part. AR is a linear combination of previous values, 'I' is an operator for differencing a time series and MA is a linear combination of previous errors. An ARMA(p, q) model of orders p and q is defined by

* This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070).

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (1)$$

where $\{\phi_i\}$ and $\{\theta_i\}$ are the parameters of the autoregressive and moving average parts respectively, and ε_t is white noise with mean zero and variance σ^2 . We assume ε_t is normally distributed, that is, $\varepsilon_t \sim N(0, \sigma^2)$. ARIMA(p, d, q) then represents the d th difference of the original series as a process containing p autoregressive and q moving average parameters. The method of building an appropriate time series forecast model is an iterative procedure that consists of the implementation of several steps. The main four steps are: identification, estimation, diagnostic checking, and forecasting. For details see [2].

B. Asymmetric ARCH time series models

Among the field of applications where the standard ARIMA fit is poor are financial and monetary problems. Exchange rates, stock market returns and other macroeconomic variables of generally high frequency are likely to originate from low complexity chaos. Detection of nonlinear hidden pattern in such time series provides important information about their behavior and improves the forecasting ability over short time. In this context, ARCH models introduced by Engle [3] arose as an appropriate framework for studying these problems. Bollerslev [4] proposed a useful extension of Engle's ARCH model known as the generalised ARCH (GARCH). In the literature several variants of basic GARCH model has been derived. In the basic GARCH model if only squared residuals ε_{t-i} enter the equation, the signs of the residuals or shocks have no effects on conditional volatility.

The basic GARCH model can be extended to allow for leverage effects. This is performed by treating the basic GARCH model as a special case of the power GARCH (PGARCH) model proposed by Ding, Granger and Engle [4].

$$\sigma_t^d = \alpha_0 + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| + \gamma_i \varepsilon_{t-i})^d + \sum_{j=1}^q \beta_j \sigma_{t-j}^d \quad (2)$$

where d is a positive exponent, and γ_i denotes the coefficient of leverage effects [4].

Commonly used asymmetric volatility models are the ARCH type models. Especially the TGARCH (threshold GARCH) and EGARCH (Exponential GARCH) models will be applied, which allows for leverage effects. TGARCH models divides the distribution of the innovations into disjunctive intervals and then approximate a piecewise linear function for the conditional standard deviation or the conditional variance respectively. However, a stylized fact of financial volatility is that bad news (negative shocks) tends to have a larger impact on volatility than good news (positive shocks). Nelson [5] proposed the following exponential GARCH model abbreviated as EGARCH to allow for leverage effects in the form

$$\log h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \frac{|\varepsilon_{t-i}| + \gamma_i \varepsilon_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^q \beta_j \log h_{t-j} \quad (3)$$

Note if ε_{t-i} is positive or there is “good news”, the total effect of ε_{t-i} is $(1 + \gamma_i) \varepsilon_{t-i}$. However contrary to the “good news”, i.e. if ε_{t-i} is negative or there is “bad news”, the total effect of ε_{t-i} is $(1 - \gamma_i) |\varepsilon_{t-i}|$. Bad news can have a larger impact on the volatility. Then the value of γ_i would be expected to be negative [6].

As we mentioned early, another extension of the classic GARCH model that allows for leverage effect is the threshold GARCH. TGARCH models divide the distribution of the innovations into disjunctive intervals and then approximate a piecewise linear function for the conditional standard deviation or the conditional variance respectively [7]. TGARCH models have therefore the following form:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i S_{t-i} \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j} \quad (4)$$

where $S_{t-i} = 1$ if $\varepsilon_{t-i} < 0$ and $S_{t-i} = 0$ if $\varepsilon_{t-i} \geq 0$.

Depending on the threshold value, ε_{t-i}^2 will have different effects on the conditional variance σ_t^2 , as it follows: when ε_{t-i} is positive, total effects are given by $\alpha_i \varepsilon_{t-i}^2$, when ε_{t-i} is negative, total effects are given by $(\alpha_i + \gamma_i) \varepsilon_{t-i}^2$.

C. Neural Approach

For the investigation with neural networks an RBF (soft, classic and granular) three layer feed-forward net is employed, where the output layer weight are trained by using backpropagation algorithm, whereas the hidden layer weights are found by a clustering algorithm applied to the input data which is an unsupervised learning technique. The transfer function in the hidden layer is a radial basic function, whereas for the output unit a linear transfer function is applied. Despite the fact that RBF neural networks possess a number of attractive properties such as the universal approximation ability and parallel structure, they still suffer from problems like the existence of many local minima and the fact that it is unclear how one should choose the number of hidden units. In order to avoid over-fitting and data-fitting the networks are kept simple (the number of hidden units in varied between 3 and 10). For more details see [8].

III. Data and Model Estimates

We illustrate the ARCH/ARCH methodology on the developing a forecast model for daily EUR/USD exchange rates time series. This time series was obtained from <http://oanda.com/currency/historical-rates/> for period from 2001 till 2010, it includes total of 3652 observations. We have 10 years long time series of the closing rates of EUR/USD exchange rates (see Fig.1).



Fig. 1 Time series of the daily exchange rates (2001 – 2010): EUR currency against the US dollar (USD).

The time series in Fig. 1 exhibits non-stationary behaviour. However, after its first differencing is stationary. To build a forecast model the sample period (Jan 2001 – Oct 2010 training data set denoted A) was defined and the ex post forecast period (Nov 2010 – Dec 2010 as validation data set denoted E).

Input selection is crucial importance to the successful development of an ARIMA-ARCH model. Tentative identification of an ARIMA time series model is done through analysis of actual historical data. The primary tools used in identification process are autocorrelation and partial autocorrelation functions (ACF, PACF). The theoretical ACF and PACF are unknown and must be estimated by the sample ACF and PACF. According to these criterion we tentatively identify the underlying model of our series to by stationary ARIMA(1, 1, 1) with the equation as follows

$$\Delta y_t = \phi_1 \Delta y_{t-1} + \phi_1 \varepsilon_{t-1} + \varepsilon_t \quad (5)$$

where Δ is the difference operator defined as $\Delta y_t = y_t - y_{t-1}$.

As we mentioned early, high frequency financial data, like our EUR/USD exchange rate time series, reflect a stylized fact of changing variance over time. An appropriate model that would account for conditional heteroscedasticity should be able to remove possible nonlinear pattern in the data. Various procedures are available to test an existence of ARCH or GARCH. A commonly used test is the LM (Lagrange Multiplier) test. The LM test assumes the null hypothesis $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_p = 0$ that there is no ARCH. The LM statistics has an asymptotic χ^2 distribution with p degrees of freedom under the null hypothesis. For calculating the LM statistics see for example [2, 5]. The LM test performed on the EUR/USD exchange rates indicates presence of autoregressive conditional heteroscedasticity.

In many cases, the basic GARCH model with normal Gaussian error distribution provides a reasonably good model for analyzing financial time series and estimating conditional volatility. However, there are some aspects of the model which can be improved so that it can better capture the characteristics and dynamics of a particular time series. For this purpose the Quantile-Quantile (QQ) plots are used. For example, the R system (<http://cran.r-project.org/>) assist in performing residual

analysis (computes the Gaussian, studentized and generalized residuals with generalized error distribution – GED) with relevant AIC criterions and Likelihood function \hat{s} values. Parameter estimates were obtained by ML (Maximum Likelihood) method using the R2.6.0 software. Our final model has the form

$$\Delta \hat{y}_t = 0.000126 \Delta y_{t-1} - 0.041943 \varepsilon_{t-1} \quad (6)$$

for mean equation, and

$$h_t = 2.46 \cdot 10^{-7} + 0.030278 y_{t-1}^2 + 0.969150 h_{t-1} \quad (7)$$

for GARCH(1,1) model with GED distributions.

Finally to test for nonlinear patterns in EUR/USD exchange rates the fitted standardized residuals $\hat{\varepsilon}_t = e_t / \sqrt{h_t}$ were subjected to the BDS test. The BDS test (at dimensions $N = 2, 3$, and tolerance distances $\varepsilon = 0.5, 1.0, 1.5, 2.0$) finds no evidence of nonlinearity in standardized residuals of the EUR/USD exchange rates time series. The fitted vs. actual EUR/USD exchange rates for the validation data set are graphically displayed in Fig. 2.

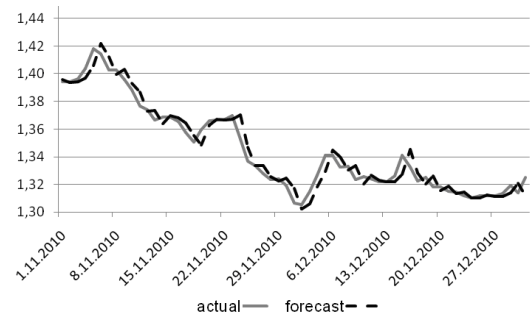


Fig. 2 Actual (solid) and forecast (dotted) values of the EUR/USD exchange rate forecast - statistical alternative, model ARIMA(1,1,1) + GARCH(1,1).

IV. Neural Approach

The same data used for estimating the linear AR(3)+EGARCH(1,1) model was also used to train neural network. The variables forming the right hand of the model given by Eq. (1) above were used as input units into the neural network.

For the investigation of the neural networks a granular RBF (Radial Basic Function) net was employed [9], (see Fig. 3), where the weights v_j have been trained by using *Backpropagation*. The transfer function in the hidden layer is a Gaussian RBF with cloud concept [9, 10], whereas for the output unit a linear transfer function is applied. The output values o_j^N from the hidden layer are “normalized”, i. e. the values whose sum is equal to 1. The weights w_j are the centers c_j of the radial basis functions [10]. To find the weights w_j or centers c_j of activation functions we used the adaptive (competitive learning) version of Kohonen's rule for clustering input data set.

As the neural network in this study is used as a non-linear supplement of linear ARIMA (1,1,1) + GARCH(1,1) approach, the network architecture allows linear input/output-links to be included.

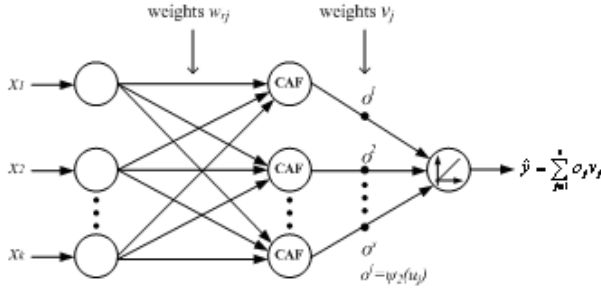


Fig. 3 Granular RBF neural network architecture with cloud activation function (CAF).

In order to avoid over-fitting, the network was kept simple (the number of hidden units were 10 the learning rate was set to 0.05. The ex post forecasts evaluation by statistical summary measures of model's forecast accuracy are given in Table I.

TABLE I Comparison of forecast summary statistics for EUR/USD exchange rate time series - statistical and neural approach: ex post period

Model	Forecast accuracy		
	RMSE	MAE	MAPE
ARIMA(1,1,1) + GARCH(1,1)	0.00793	0.00646	0.00495
Neural Approach Inputs: $\Delta y_{t-1}, \varepsilon_{t-1}$	0.00185	0.00145	0.00107

The fitted vs. actual EUR/USD exchange rates for the validation data set are graphically displayed in Fig. 4.

V. Empirical Comparison and Discussion

From Table I it is shown that both forecasting models used are very accurate. The development of the error rates on the validation data set showed a high inherent deterministic relationship of the underlying variables. Though promising results have been achieved with both approaches, for the chaotic financial markets a purely linear (statistical) approach for modeling relationships does not reflect the reality. For example if investors do not react to a small change in exchange rate at the first instance, but after crossing a certain interval or threshold react all the more, then a non-linear relationship between Δy_t and $\Delta y_{t-3}, \varepsilon_{t-1}$ exist in model (6).

The training process and development of neural approach based on G RBF NN not only detected the functionality between the underlying variables as well as the short-run dynamics. Moreover, as we could see, the RBF NNs have such attributes as computational efficiency, simplicity, and ease adjusting to changes in the process being forecast. Thus,

neural networks are usually used in the complicated problems of prediction because they minimize the analysis and modeling stages and the resolution time. Thus, we can expect more interests of the people who will make and at the same time use the forecast. If the managers are convinced that the forecasting system is sound and , they may make little use of the information given to them.

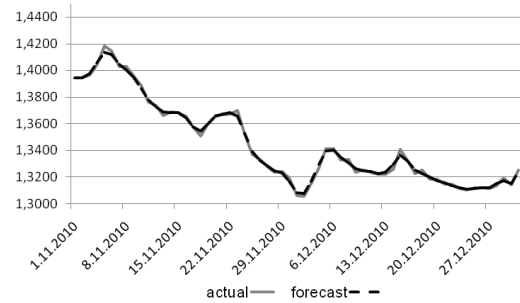


Fig. 4 Actual (solid) and forecast (dotted) values of the EUR/USD exchange rate forecast (neural approach).

VI. Conclusion

The results of the study showed that there are more ways of approaching the issue of risk reducing in managerial decision-making in companies, financial institutions and small enterprises. It was also proved that it is possible to achieve significant risk reduction in managerial decision-making by applying modern forecasting models based on information technology such as neural networks developed within artificial intelligence. In future research we plan to extend presented methodologies by applying fuzzy logic systems to incorporate structured human knowledge into workable learning algorithms.

References

- [1] CH. Gouri éroux, *ARCH Models and Financial Application*. Springer Verlag, New York, 1997.
- [2] G.E.P. Box and G.M. Jenkins, *Time Series Analysis, Forecasting and Control*. Revised Edition, Holden-Day, San Francisco, CA, 1976.
- [3] R.F. Engle, Auto Regressive Conditional Heteroscedasticity with Estimates of the Variance of United Kindom Inflation. *Econometrica*, 50 (1982), pp. 987-1007.
- [4] D. Bollerslev, (1986). Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, 31, pp. 307-327, 1986.
- [5] Z. Ding, C.W. Granger, and R.F. Engle, A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance*, 1, pp. 83-106, 1993.
- [6] E. Zivot, J. Wang, *Modeling Financial Time Series with S-PLUS®*. Springer Verlag, NY, 2005.
- [7] J.M. Zakoian, Treshold Heteroscedastic Models, *Journal of Economic Dynamics and Control*, 18, pp. 931-955, 1994.
- [8] V. Kecman, *Learning and soft computing: support vector machines, neural networks, and fuzzy logic*. Massachusetts: The MIT Press, 2001.
- [9] D. Li, and Y. Du, *Artificial intelligence with uncertainty*. (Boca Raton: Chapman & Hall/CRC, Taylor & Francis Group, 2008.
- [10] M. Marcek D. Marcek, Granular RBF Neural Network Implementation of Fuzzy Systems: Application to Time Series Modelling, *Journal of Mult.-Valued Logic & Soft Computing*, Vol. 14, No. 3-5, pp. 101-114, 2008.