GARCH MODEL INDENTIFICATION USING NEURAL NETWORK

DSc. André Machado Caldeira
Fuzzy Consultoria, Brazil
E-mail: mmachado@ibmecrj.br

DSc. Maria Augusta Soares Machado
IBMEC/RJ, Brazil
E-mail: mmachado@ibmecrj.br

Ph.D. Reinaldo Castro Souza
Pontificia Universidade Católica (PUC/RJ), Brazil
E-mail: reinaldo@ele.puc-rio.br

Ph.D. Ricardo Tanscheit
Pontificia Universidade Católica (PUC/RJ), Brazil
E-mail: ricardo@ele.puc-rio.br

ABSTRACT

GARCH models are being largely used to estimate the volatility of financial assets, and GARCH (1,1) is the one most used. However, identification of GARCH models is not fully explored. Some specialist systems technology have been used in some applications of time series models such as time series classification problems, ARMA models identification, as well as SARIMA. The aim of this paper is to develop an intelligent system that can accurately identify the specification of GARCH models providing the right choice of the model to be used, thus avoiding the indiscriminate usage of GARCH (1,1) model.

Keywords: GARCH, Volatility, Identification.

527

1   INTRODUCTION

"Identification of the right GARCH model specification, to be adjusted for a time series, is generally difficult. So it is recommended to use low orders models, like (1,1), (1,2) or (2,1), and then choose the best one using a criteria, for example AIC or BIC, ..." (MORETTIN; TOLOI, 2004).

ARCH and GARCH models have being largely explored technically and empirically since their creation in 1982 and 1986, respectively. However, the focus is always on stylized facts of financial time series or volatility forecast, where GARCH (1,1) is commonly used. Hardly ever do we find a study concerning the identification of GARCH models. Some studies have been developed using specialized systems applied to time series models (REYNOLDS, et al., 1995) and identification of both ARMA (MACHADO, 2000) and SARIMA (SILVA, 2005) models. In this context, this paper has as its aim the development of an intelligent system which could improve the specification identification, thus avoiding the indiscriminate usage of GARCH (1,1) model. In order to validate the accuracy and efficacy of the system proposed, simulated time series will be used. The results derived from such system will then be compared to chosen model derived from AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) criteria.

This paper is developed in five chapters. The following chapter presents theoretical concepts relevant to this paper such as the foundation to the development of the system. The third and fourth chapters present the identification results using AIC and BIC criteria and using the specified system proposed by this paper. The concluding chapter focuses on further discussion on the subject meter and proposes new development.

1.1   GARCH Models

Currently, financial markets suffer significant influence of daily news. On analyzing a series of financial asset returns which present a shift between periods of high and low volatility forming clusters, volatility can be defined "as a conditional variance of a time series" (VEIGA, et al., 1993).

During the high-level volatility period, the investor may feel reluctant to invest, and as a consequence many assets values are penalized because of their liquidity. However, when volatility is not so high, it is good for the financial market.

The excess of volatility can bring many consequences into the financial market such as:

- in asset prices: the volume of investment reduces and investors are induced to change from a high-risk to a low-risk asset in other markets;

- in interest rate: the cost of credit increases, and as a consequence there may be an impact on the economy level;

- in currency exchange rate: whenever there is a significant decrease in the total amount of importation, the price of important and exported goods may increase due to exchange rate risks. In addition there may be a decrease in consumption levels of imported goods. .

Volatility is extremely important for the economy and financial markets, and by taking this into account, studies concerning financial time series are being developed using models different then the classic ARMA time series models (BOX; JENKINS, 1976). Such classic models cannot reproduce financial time series with essentials characteristics known as stylized facts.

Many kinds of models have being developed to estimate volatility, for example, the Exponential Weight Moving Average (Known as Risk Metrics), stochastic volatility models and GARCH models. This study focuses on GARCH models; for further details about other models see Clark (1973), Taylor (1980, 1986 e 1994), Tauchen and Pitts (1983), Hull and White (1987) and Harvey et al. (1994).

The concepts of stylized facts of financial time series are really necessary to understand the inspirations of GARCH models. For further information on stylized facts of financial time series, see Bernardo and Fernandes (1999).

The main stylized facts of financial time series could be ranked as such:

- **stylized fact 1**: Stationary Series – Statistical proprieties are static over time.

- **stylized fact2**: Weak or no linear dependence and non-linear dependence (GARCH effect). Series are not or are little auto-correlated, but the quadratic series are auto-correlated.

- **stylized fact3**: Non-Gaussian – Financial time series commonly presents skewness and higher kurtosis.

- **stylized fact4**: Existence of volatility clusters – Financial time series commonly present alternate periods of high volatility and low volatility. The conditional variance is time dependent.

A central hypothesis of the option valuation model proposed by Black e Scholes (1973) is that the financial time series performs as a Brownian movement, or the distribution of the returns is log-normal with the same mean and variance over the time. However, Mandelbrot (1963) and Fama (1963 and 1965) proposed that those series have higher kurtosis and they discussed the existence of volatility clusters. Those characteristics were interpreted as an evidence of stochastic volatility of financial assets.

For the purpose of representing those characteristics, since approximately two decades ago, GARCH models are being largely used in financial studies, especially in financial derivatives studies. The initial success of ARCH models to represent the non-linear dependence made possible many extensions.

1.1.1 GARCH models representation

The first model from the GARCH family was introduced by Engle (1982). This model can represent some stylized facts of a financial time series. Engle proposed to model the quadratic of the return time series using an autoregressive model with q parameters (AR(q)). This model was called Autoregressive Conditional Heteroskedastic or ARCH(q), which can be written by the expression:

$$u_t^2 = \zeta + a_1 u_{t-1}^2 + a_2 u_{t-2}^2 + \ldots + a_q u_{t-q}^2 + w_t \qquad (1.1)$$

Where w_t is a white noise:

$$
\begin{aligned}
&E(w_t) = 0 \\
&E(w_t w_z) = \begin{cases} \lambda^2 & for\ t = \tau \\ 0 & in\ other\ cases \end{cases}
\end{aligned}
\qquad (1.2)
$$

Sometimes it is convenient to re-write this expression like this:

Suppose:

$$u_t = \sqrt{h_t}\, \upsilon_t \qquad (1.3)$$

Where:

$$E(\upsilon_t) = 0 \qquad E(\upsilon_t^2) = 1 \qquad (1.4)$$

If h_t is written as:

$$h_t = \zeta + a_1 u_{t-1}^2 + a_2 u_{t-2}^2 + \ldots + a_q u_{t-q}^2 \qquad (1.5)$$

This implies:

$$E\left(u_t^2 \mid u_{t-1}, u_{t-2}, \ldots\right) = \zeta + a_1 u_{t-1}^2 + a_2 u_{t-2}^2 + \ldots + a_q u_{t-q}^2 \qquad (1.6)$$

So, if u_t is generated by (1.3) e (1.5), then u_t follows an ARCH(q) process, and if (1.3) and (1.5) are used in (1.1), it becomes:

$$h_t \upsilon_t = h_t + w_t \qquad (1.7)$$

Using (1.3) specification, the w_t innovation in AR(q) representation for u_t^2 in (1.1) can be expressed by:

$$w_t = h_t\left(\upsilon_t^2 - 1\right) \qquad (1.8)$$

Notice that even if the unconditional variance of w_t is assumed to be a constant in (1.8), the conditional variance of w_t changes over time. Thus, the ARCH model can describe volatility clusters.

In 1986, Bollerslev observed, by empirical evidence, that it would be necessary to estimate ARCH models with high orders to reproduce the conditional variance dynamics. In order to solve this problem, he proposed a more general and parsimonious form of ARCH model, which he called Generalized Autoregressive Conditional Heteroskedastic (GARCH) (BOLLERSLEV, 1986).

The same idea of parsimony used in ARMA models was then applied to GARCH models. So, it can be demonstrated that a Moving Average model (MA) with

order one is equivalent to an Autorregressive model (AR) with infinite order. In order to reduce the number of parameters to be used, the AR is merged with MA, thus creating the ARMA model. GARCH model is based on ARCH model with infinite order and h_t can be expressed as:

$$h_t = k + \sum_{i=1}^{q} \alpha_{t-i} u_{t-i}^2 + \sum_{j=1}^{p} \beta_{t-j} h_{t-j}^2 \tag{1.9}$$

For the same reason that ARCH models depend on some restrictions concerning h_t to be positive for every t, GARCH models depend on restrictions of $k > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$. Nelson and Cao (1992) observed that the conditions $\alpha_i \geq 0$ and $\beta_j \geq 0$ were sufficient, but not necessary. So, they argued that by imposing such conditions could be excess of precaution and could become a limitation considering some empirical works, and in practical applications, even if there is some negative coefficients, the conditional variance becomes positive. Such restrictions could be relaxed and in practical works it is used to estimate the coefficients with none of those restrictions.

In many applications using high frequency time series, the estimated conditional variance by a GARCH (p,q) model demonstrates a strong persistence, that is:

$$\sum_{j=1}^{p} \beta_j + \sum_{i=1}^{q} \alpha_i \approx 1$$

If $\sum_{j=1}^{p} \beta_j + \sum_{i=1}^{q} \alpha_i < 1$, the process ($u_t$) is second order stationary and the noise on the conditional variance of σ_t^2 has a decrease impact on σ_{t+h}^2, when h increases, and is asymptotically insignificant. This feature is called persistence.

Other variations of GARCH models were proposed having in mind many objectives, as for example the Exponential GARCH (EGARCH) (NELSON, 1991; ENGLE; NG, 1993) and the TGARCH (ZAKOIAN, 1991; GLOSTEN, et al., 1993; RABEMANANJARA; ZAKOIAN, 1993), that were proposed to capture the asymmetric effect on the volatility clusters

532

GARCH and ARCH models will be applied in this paper.

1.1.2 Modeling Strategy

Franses and Djik (2000) proposed a modeling sequence which uses the following steps:

- calculate some time series statistics (ACF, Auto-correlation Function, and PACF, Partial Auto-correlation Function);

- compare those values with theoretical values to specify the right model (Identification);

- estimate parameters of the specified model (Estimation);

- evaluate the specified model using adequacy metrics (Validation);

- re-specify the model if necessary;

- use the model to make the forecast (Forecasting).

The specification of the appropriate structure (identification) for the equation of the conditional variance of a time series which follows a GARCH process is the main concern of this paper. Autocorrelation function (ACF) and Partial Autocorrelation function (PACF) are commonly used in the identification and validation of the ARMA model specification (BOX; JENKINS, 1976). On the other hand, Bollerslev et al. (1988) showed that those functions, when applied on the square of the time series, could be used to the specification and validation of the GARCH model.

Suppose that τ_n is the n-th autocorrelation and ϕ_{kk} is the k-th partial autocorrelation of u_t^2 obtained through the solution to the equations for the GARCH models, analogues of Yule-Walker equations. Thus, the usual interpretation for ARMA models can be used for GARCH models. For an ARCH(q) process, ϕ_{kk} has an abrupt cut after the q-lag, which behavior is identical to the partial autocorrelation function of an AR(q) process. On the other hand, the autocorrelation function of u_t^2 for a GARCH(q,p) process is different from zero and has an exponential decay. By using these patterns, such functions can help identify the right specification of the GARCH model.

Another way to identify the specification of GARCH models is to use the AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) statistics. The model which shows the lowest statistic is the one selected to be the identified model. Some results using this means of identification are presented in chapter 3.

2   PROPOSED IDENTIFICATION METHODOLOGY

The proposed identification methodology blend the procedure of autocorrelation and partial autocorrelation functions described by Bollerslev, et al. (1988) with the identification using AIC and BIC, and further test over-specification of Box and Jenkins (1974). **Erro! Fonte de referência não encontrada.** represents such methodology.

First step is to train a neural network to represent the pattern configuration of each model using autocorrelation, partial autocorrelation function and statistics AIC and BIC. Therefore, the next step is to test models with high orders then the selected one. Using both steps the final identification is done.
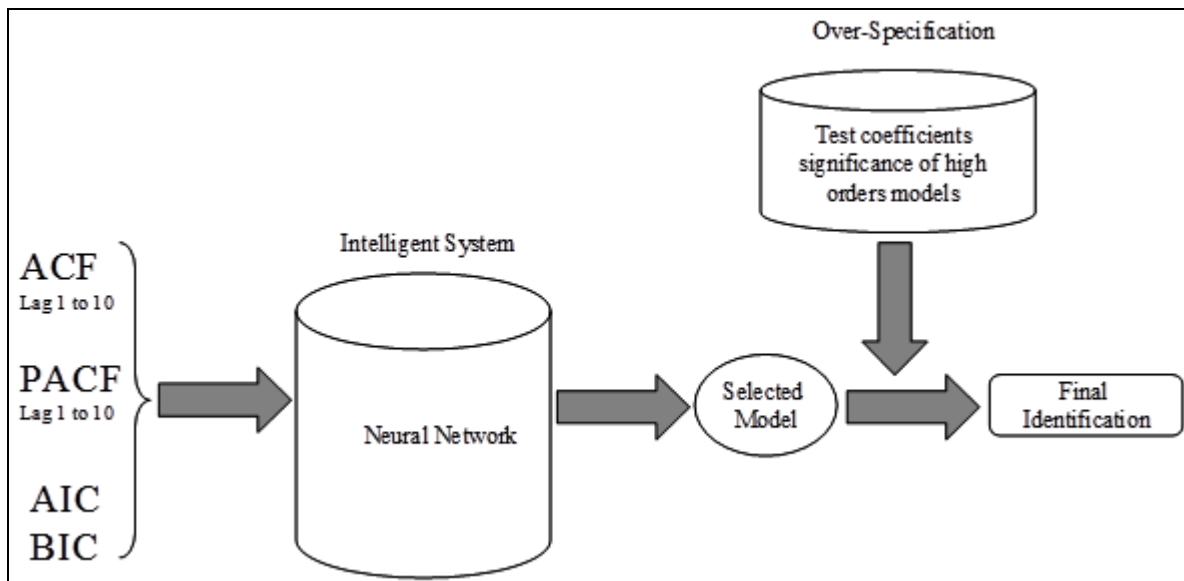

Figure 1: Proposed Identification Methodology

3   Applied Study on Simulated Data

In order to compare the identification performance of the statistics AIC and BIC with the proposed neural network, the first step is to simulate a time series sample generated by GARCH processes using MatLab software for this purpose.

The models to be compared are ARCH(1), ARCH(2), GARCH(1,1), GARCH(2,1) and GARCH(1,2). The simulated data total showed 8,000 series, of

which 1,600 were series generated for each model, divided in four lengths of series, in which one moth was represented by 22 observations, one quarterly period was represented by 66 observations, one semester was represented by 132 observations and one year was represented by 264 observations. Each length had 400 series for each model.

Random numbers between zero and one were used to represent the coefficients of the specified model, taking into account two restrictions: lower lags have higher coefficient than higher ones, and the sum of all coefficients is lower than one, which is a condition for GARCH models.

3.1   Model Identification using AIC e BIC

By using those simulated data, the model selected as the best was the one which has the lower AIC and BIC. Table 1 show this identification criteria results. Because the data are simulated, the generated model is known. So, it is possible to know whether AIC or BIC classified them with accuracy or not.

Table 1: Results of identification using AIC and BIC

Series Length observations	Correctly classified series by AIC identification		Correctly classified series by BIC identification	
	Series	Percentage	Series	Percentage
22	488	24.4%	465	23.3%
66	809	40.5%	734	36.7%
132	1,070	53.5%	947	47.4%
264	1,371	68.6%	1,200	60.0%
Total	3,738	46.7%	3,346	41.8%

Identification with AIC and BIC present high level of misclassified percentage, higher than 50% considering the total data classification. Considering just the annual series, (264 observations) that identification reached almost 70% of correctly classified series, but taking a look on the smaller series, the results presented a lower level of correctly classified series. For example, considering data from recent Initial Public Offering (IPO), those data should probably show high probability of misclassification.

Tables 2 and 3 show the percentage of correctly classified series of each model using AIC and BIC criteria. It can be observed that when the number of parameters increases, misclassification also increases, as it is already expected by the time that AIC and BIC penalize the model when a new parameter is introduced

with the aim of looking for parsimony. Therefore, those criteria tend to bias the classification due to parsimony.

As long as the AIC presents a higher percentage of correctly classified series, such criteria will be used from now on as a benchmark in this study.

Table 2: Percentage of correctly classified series using AIC

Series Length observations	ARCH(1)	ARCH(2)	GARCH(1,1)	GARCH(2,1)	GARCH(1,2)	Total
22	94.5%	22.5%	3.5%	1.5%	0.0%	24.4%
66	92.8%	64.8%	29.5%	8.0%	7.3%	40.5%
132	92.3%	82.5%	54.0%	23.3%	15.5%	53.5%
264	93.0%	89.8%	74.5%	53.0%	32.5%	68.6%
Total	93.1%	64.9%	40.4%	21.4%	13.8%	46.7%

Table 3: Percentage of correctly classified series using BIC

Series Length observations	ARCH(1)	ARCH(2)	GARCH(1,1)	GARCH(2,1)	GARCH(1,2)	Total
22	96.8%	16.0%	3.0%	0.5%	0.0%	23.3%
66	98.8%	56.5%	25.0%	2.0%	1.3%	36.7%
132	99.3%	79.0%	53.8%	3.5%	1.3%	47.4%
264	99.5%	91.5%	83.0%	19.8%	6.3%	60.0%
Total	98.6%	60.8%	41.2%	6.4%	2.2%	41.8%

3.2   Intelligent System Identification

As presented in section 2, the first step of intelligent system identification is to specify the neural network to be trained. Figure 2 represents proposed neural network specification.
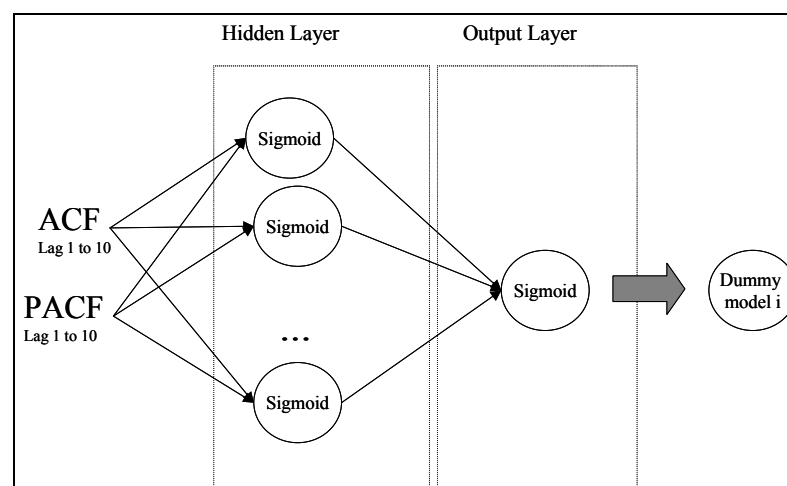

Figure 2: Neural Network Specification to identify GARCH structure

Neurons of the hidden layers and the neuron of the output layer are represented by sigmoid functions. Once more MatLab software was used.

The same 8,000 series of section 3.1 were used to test the neural network. For the training data, were generated other 8,000 series with the same characteristics of the simulated data as seen in section 3.1.

Before specifying the best structure for the neural network to be trained, it is needed to select which input features will be used, therefore it was applied the Fischer Score feature selection method (BISHOP, 1995). Such method selected the following variables: ACF (lag1), ACF (lag2), ACF (lag3), PACF (lag1), PACF (lag2), Difference between ACF (lag2) and ACF (lag1), and Difference between ACF (lag3) and ACF (lag2).

After the feature selection, neural network topology needs to be specified, so a sensitive analysis was done, varying the number of neurons and the number of layers, the results can be seen on Figure 3.
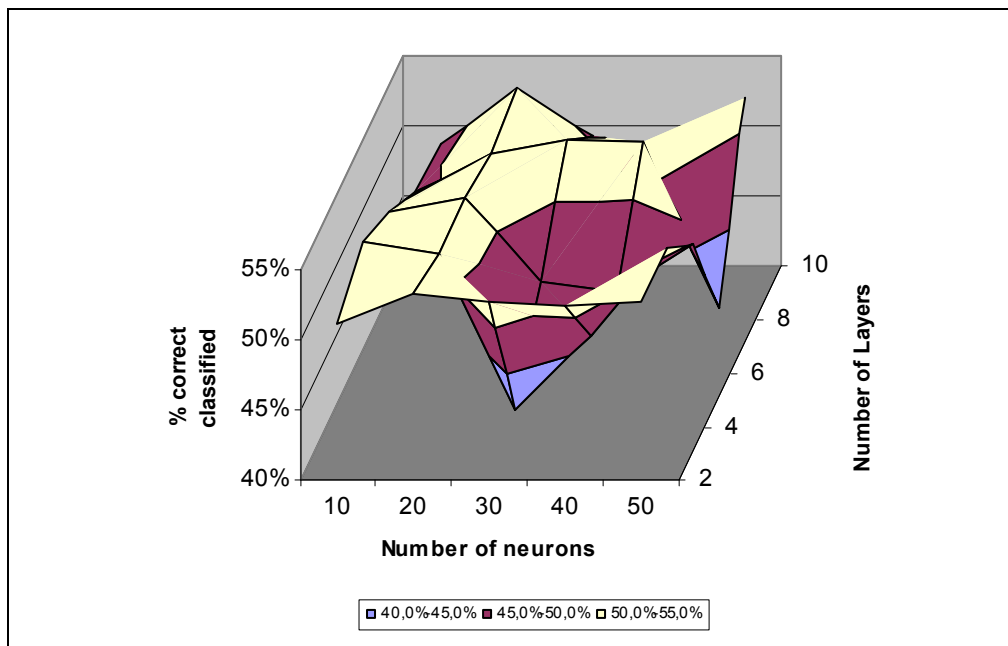

Figure 3: Accuracy by varying number of layer and number of neurons

By analyzing previews chart it can be seen that the best result was reached by the topology with two hidden layers using twenty neurons in each hidden layer. It can also be observed that the misclassification of the neural network with lower layers increases as the number on neurons increases. Such results might indicate over-fitting.

Table 3 and Table 4 present classification results using neural network methodology.

Table 3: Percentage of correctly classified series using Neural Network

Series Length observations	ARCH(1)	ARCH(2)	GARCH(1,1)	GARCH(2,1)	GARCH(1,2)	Total
22	66.3%	23.5%	27.3%	9.0%	38.8%	33.0%
66	86.0%	61.8%	25.5%	8.5%	54.5%	47.3%
132	89.3%	81.5%	42.5%	23.5%	62.8%	59.9%
264	92.0%	89.5%	57.0%	53.0%	69.3%	72.2%
Total	83.4%	64.0%	38.1%	23.5%	56.3%	53.1%

Table 4: Cross-classification percentage using Neural Network

Real / Classified	ARCH(1)	ARCH(2)	GARCH(1,1)	GARCH(2,1)	GARCH(1,2)
ARCH(1)	**84.1%**	2.9%	3.1%	0.8%	9.1%
ARCH(2)	15.8%	**64.0%**	5.9%	6.8%	7.6%
GARCH(1,1)	21.1%	15.4%	**33.3%**	6.6%	23.6%
GARCH(2,1)	12.8%	26.7%	17.3%	**29.0%**	14.2%
GARCH(1,2)	17.5%	4.7%	16.0%	4.9%	**56.9%**

The experiment suggests that AIC and BIC can be improved by using computational intelligence. The specified neural network presented 53.1% of correctly classified series, representing an improvement of 640 bps considering the results of AIC, and if they are compared to the BIC results, there is an improvement of 1130 bps, especially considering GARH(1,2).

Those results suggest that the neural network methodology increases the percentage of correctly classified series. However, Table 4 shows many GARCH(2,1) misclassified as GARCH(1,1) or ARCH(2), and GARCH(1,1) misclassified as ARCH(1), for example. Even though, it is notorious that ARCH(1) and ARCH(2) have better classifications results.

Those results can be improved by over specifying such models. In other words, the number of parameters of the identified model can be increased and its significance tested. Thus, for this purpose, Table 5 describes which models are tested using the T-test for the significance of the new parameter.

Table 5: Overspecify procedure

Identified Model	Overspecified model 1	Overspecified model 2
ARCH(1)	ARCH(2)	GARCH(1,1)
ARCH(2)	GARCH(2,1)	None
GARCH(1,1)	GARCH(2,1)	GARCH(1,2)
GARCH(2,1)	None	None
GARCH(1,2)	None	None

Figure 4 shows the classification performance as the significance of T-test is applied on the over specifying procedure.
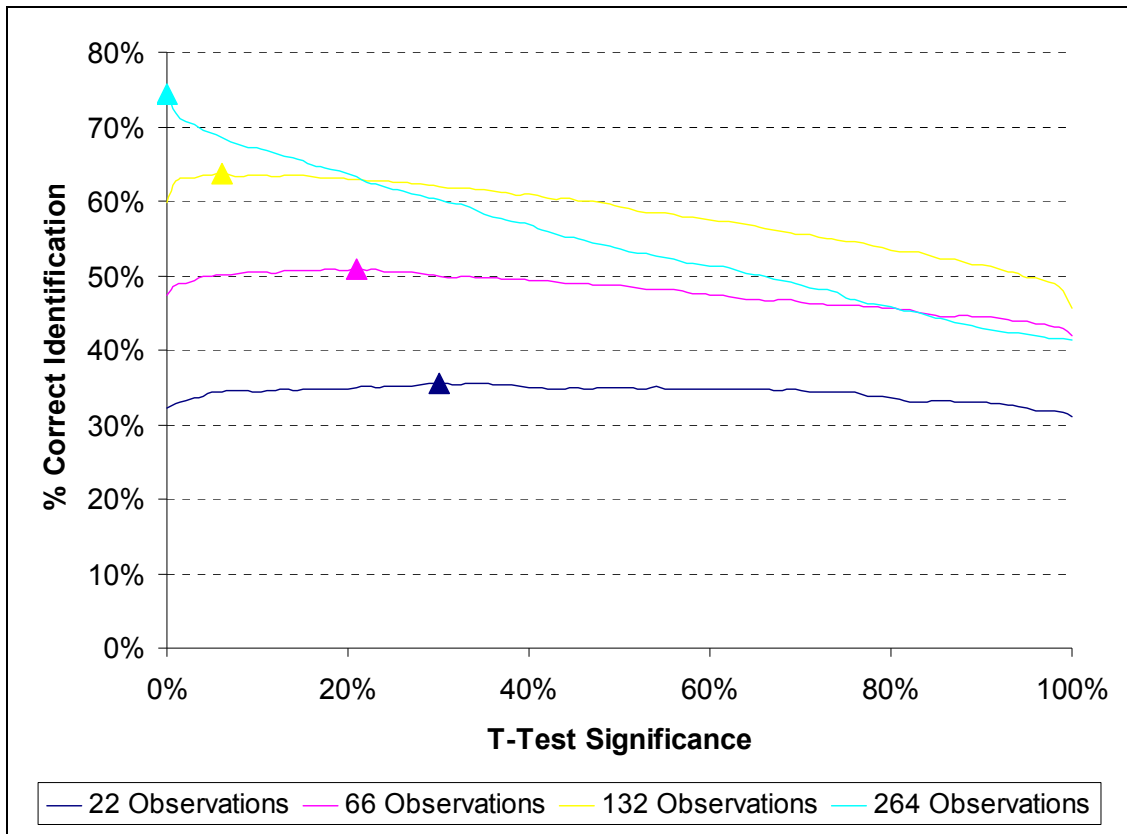
Figure 4: Accuracy by varying T-test significance

By applying best results for each length of series, the performance improved by almost 5%. The results can be shown on Table 6 and Table 7.

Table 6: Right classification percentage using Neural Network after over specifying

N	ARCH(1)	ARCH(2)	GARCH(1,1)	GARCH(2,1)	GARCH(1,2)	Total
22	55.8%	24.5%	26.0%	23.3%	48.5%	**35.6%**
66	72.0%	57.5%	35.3%	41.8%	47.8%	**50.9%**
132	86.3%	79.8%	44.8%	50.5%	56.8%	**63.6%**
264	92.5%	91.0%	56.8%	54.3%	77.3%	**74.4%**
Total	**76.6%**	**63.2%**	**40.7%**	**42.4%**	**57.6%**	**56.1%**

Table 7: Cross-classification percentage using Neural Network after over specifying

Real / Classified	ARCH(1)	ARCH(2)	GARCH(1,1)	GARCH(2,1)	GARCH(1,2)
ARCH(1)	**76.6%**	5.9%	7.4%	0.9%	9.1%
ARCH(2)	13.8%	**63.2%**	6.4%	8.6%	8.0%
GARCH(1,1)	8.2%	16.3%	**40.7%**	10.3%	24.5%
GARCH(2,1)	1.4%	18.5%	21.7%	**42.4%**	16.0%
GARCH(1,2)	0.4%	7.3%	23.1%	11.7%	**57.6%**

Table 6 shows that the over specification improved the results by almost 5%, so if those results are compared to the AIC results they present an increase of 20% of right identification on overall results. They improved the overall performance classification from 47% of the AIC identification to 56% of neural network.

4   FINAL CONSIDERATIONS

As presented in section 3.1, the statistics AIC and BIC were able to classify correctly just 46.7% and 41.8% respectively. However, if the annual series are excluded, the performance reaches 39.5% and 35.8% respectively. Therefore, the performance of the proposed neural network improves considerably, from 46.7% of overall correctly identified series by the AIC to 56.1% of overall correctly identified by neural network, demonstrating that there are opportunities to gain performance in the identification of the GARH model.

As a follow up to this study, an application using real time series can be done to test predicted performance of each model selected by the criteria tested. This application can be of great importance especially to emerging capital markets as they can be good resources capitalization option to middle-sized and large-sized companies.

REFERENCES

BERNARDO, M. R.; FERNANDES, C. A. (1998) **Utilização de Modelos não-lineares não-Gaussianos para Estimação de Volatilidade de Séries Temporais Financeiras**. Dissertação (Graduação) – Pontifícia Universidade Católica do Rio de Janeiro.

BLACK F.; SCHOLES, M. (1973) The pricing of options and corporate liabilities. **Journal of Political Economy**, n. 81, p. 637-59.

BOLLERSLEV, T. (1986) Generalized autorregressive conditional heteroskedasticity, **Journal of Econometrics**, n. 31, p. 303-327.

BOLLERSLEV, T.; ENGLE, R.; WOOLDRIDGE, J. (1988) A capital asset pricing model with time varying covariances. **Journal of Political Economy**, n. 96, p. 116-131.

BOX, G. E. P.; JENKINS, G. M. (1976). **Time Series Analysis: Forecasting and Control**. Holden-Day, San Francisco.

CLARK, P. K. (1973) A subordinated Stochastic Process model with finite variance for speculative prices. **Econometrica**, n. 41, p. 135-155.

ENGLE, R. F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. **Econometrica**, v. 50, n 4, p. 987-1007.

ENGLE, R. F.; NG, V. K. (1993) Measuring and testing the impact of news on volatility. **Journal of Finance**, n. 48, p. 1749-1778.

FAMA, E. (1963) Mandelbrot and the Stable Paretian hypothesis. **Journal of Business**, n. 36, p. 420-429.

FAMA, E. (1965) The behavior of stock prices, **Journal of Business***,* n. 47, p. 244-280.

FRANSES, P. H.; DIJK, D. V. (2000) **Nonlinear time series models in empirical finance**. Cambridge University Press.

GLOSTEN, L.; JAGANNATHAN, R.; RUNKLE, D. (1993) Relationship between the expected value and the volatility of the nominal excess return on stocks. **Journal of Finance***,* n. 48, p. 1779-1801.

HARVEY, A.C.; RUIZ, E.; SHEPHARD, N. (1994) Modeling Stochastic variance models. **Review of Economic Studies**, n. 61, p. 247-267.

HAYKIN, S. (1999) **Neural networks**: a comprehensive foundation. Prentice-Hall.

HULL, J.; WHITE, W. (1987) The pricing of options on assets with stochastic volatility. **Journal of Finance**, n. 42, p. 281-300.

MACHADO, M. A. S (2000) Auxílio à Idenficação de Modelo Box & Jenkins Usando Redes Neurais Nebulosas. **Pesquisa Naval**, n. 7, p. 49.

MANDELBROT, B., (1963) The variation of certain speculative prices. **Journal of Business***,* n. 36, p. 394-419.

MORETTIN, P. A.; TOLOI, C. M. C. (2004) **Análise de séries temporais**. Edgard Blücher, São Paulo.

NELSON, D. (1991) Conditional heteroskedasticity in assets returns: a new approach. **Econometrica**, v. 59, n. 2, p. 347-370.

NELSON, D. B.; CAO, C. Q. (1992) Inequality constraints in the univariate GARCH model. **Journal of Business and Economic Statistics**, n. 10, p. 229-235.

RABEMANANJARA, R.; ZAKOIAN, J. M. (1993) Threshold ARCH models and asymmetries in volatility, **Journal of Applied Econometrics***,* n. 8, p. 31-49.

REYNOLDS, B.; STEVENS, T.; MELLICHAMP, R.; SMITH, M. J. (1995) Box-Jenkins Forecast Model Identification, **A.I. Expert**.

SILVA, L. M. (2005) **Uma aplicação de Árvores de Decisão, Redes Neurais e KNN para a Identificação de Modelos ARMA não Sazonais e Sazonais**. Tese de Doutorado em Engenharia Elétrica – Pontifícia Universidade Católica do Rio de Janeiro.

TAUCHEN, G. E.; PITTS, M. (1983) The price variability-volume relationship on speculative markets. **Econometrica**, n. 51, p. 485-505.

TAYLOR, S. J. (1980) Conjectured models for trend in financial prices, tests and forecast. **Journal of the Royal Statistical Society**, Series A, n. 143, p. 338-362.

TAYLOR, S. J. (1986) **Modeling Financial Time Series**. New York: John Wiley.

Taylor, S. J. (1994) Modeling Stochastic Volatility. **Mathematical Finance**, n. 4, p. 183-204.

VEIGA FILHO, A. L.; FERNANDES, C. A. C.; BAIDYA, T. (1993) Medidas de Volatilidade para Opções, **XXV SBPO/SOBRAPO***,* n. 1, p. 185-187.

ZAKOIAN, J. M. (1991) Threshold heteroskedastic models, **Technical report**, INSEE.