https://doi.org/10.26552/com.C.2003.4.55-58

Dušan Marček – Michaela Ačová *

ECONOMIC TIME SERIES FORECASTING: BOX-JENKINS METHODOLOGY, SIGNAL PROCESSING AND NEURAL NETWORK APPROACH

Abstract: This paper is devoted to the presentation of methods of economic time series analysis and modelling using the Box-Jenkins methodology, the signal processing approach and the feedforward neural network technique. Some results of our research on time series modelling with emphasis on potential improving forecast accuracy are presented here. The assessment of the particular models has been made using the root mean square error.

1. Introduction

Artificial neural network (ANN) is now being applied to many problems far removed from their first beginnings. Application to management problems have included predicting bankruptcy [3], predicting ratings of corporate bonds [11], forecasting financial markets [6] and time series forecasting [7]. Their main strengths lie in pattern recognition and have been a hot topic of research for many years now.

There is much controversy about the application of traditional statistical or econometric models and the ANN approaches within the field of economic time series modelling and forecasting. These controversies are based on the assumptions that there is no consensus at all on whether there is chaos in economic time series or not. Various tests for nonlinear pattern and chaos in time series have been proposed to illustrate the nonlinear nature of certain processes. A survey of these tests is presented in [1].

The goal of this paper is to illustrate those three areas: probabilistic, adaptive signal processing and computational networks may be used to economic time series modelling. In Section 2, we can see that a random process of time series of stock prices may be generated as the output of linear filter driven by white noise. Section 3 is focused on the behaviour of the GL filter and LSL when they are used to forecast future observations of stationary AR processes. In Section 4 of this paper, we report on an ANN application that was designed and run by [9] to investigate the problem of forecast accuracy across proposed models.

2. Application of the Box-Jenkins methodology in the stock prediction problem

In this section, we give an example that provides one kind of possible results. We will regard these results as referential values for the approach of adaptive signal processing procedures and ANN modelling. Many of modelling techniques of autoregressive processes are based on recent developments in time series analysis recently consolidated and presented by Box and Jenkins [4].

To illustrate the Box-Jenkins methodology, consider the stock price time readings of a typical company (say VAHOSTAV company). We would like to develop a time series model for this process so that a predictor for the process output can be developed. The data was collected for the period January 2, 1997 to December 31, 1997, which provided a total of 163 observations (see Fig. 1).

To build a forecast model the sample period for analysis y_1, ..., y_{128} was defined, i.e. the period over which the forecasting model was developed and the ex post forecast period (validation data set), y_{129}, ..., y_{163} as the time period from the first observation after the end of the sample period to the most recent observation. By using only the actual and forecast values within the ex post forecasting period only, the accuracy of the model can be calculated.

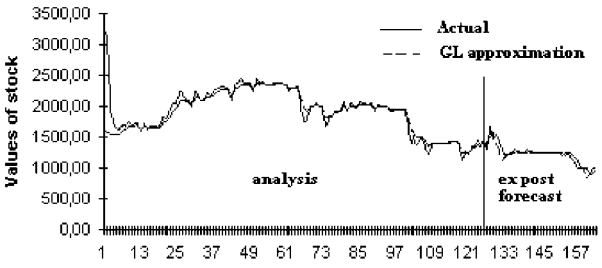

Fig. 1 The data for VAHOSTAV stock prices (January 1997 - December 1997) and the values of the AR(7) model for VAHOSTAV stock prices estimated by GL algorithm

To determine appropriate Box-Jenkins model, a tentative ARMA model in identification step is identified. In Figure 2, the estimate

* [1]Dušan Marček, [2]Michaela Ačová
[1]Faculty of Management Science and Informatics, University of Zilina, E-mail: marcek@fria.utc.sk
[2]Institute of Computer Science, Silesian University Opava, E-mail: michaela.acova@fpf.slu.cz

of autocorrelation ($\hat{r}_k$) and partial autocorrelation $\hat{a}_{kk}$ function (ACF, PACF) of the data are given. To test whether the autocorrelation and partial autocorrelation coefficients are statistically equal to zero, we use the t-statistic $t_r = \hat{r}_k / S(\hat{r}_k)$ and $t_a = \hat{a}_{kk} / S(\hat{a}_{kk})$ where

$$S(\hat{r}_k) = N^{-1/2}\left[1 + 2\sum_{j=1}^{k-1}\hat{r}_j\right]$$

and

$$S(\hat{a}_{kk}) = N^{-1/2}$$

denote standard errors of the kth sample autocorrelation or partial autocorrelation coefficient, respectively, N is the number of data points, k is the lag. Since the ACF decays in an exponential fashion, and the PACF truncates abruptly after lag 2, we may tentatively identify the model for this time series as AR(2).

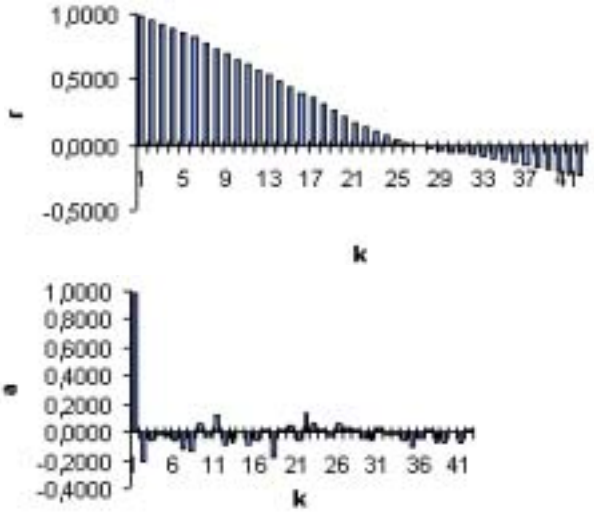

Fig. 2 Autocorrelation function and partial autocorrelation function of the data for VAHOSTAV stock prices (period for analysis)

In the estimation step, we compute estimates for the parameters of the AR(2) model

$$y_1 = \xi + a_1 y_{t-1} + a_2 y_{t-2} + \epsilon_t \quad t = 1, 2, ..., N-2 \tag{1}$$

or with of obvious matrix notation

$$y = Xa + \epsilon$$

by OLS (Ordinary Least Squared)

$$\hat{a} = (X'X)^{-1}X'y = \begin{bmatrix}\hat{\xi}\\ \hat{a}_1\\ \hat{a}_2\end{bmatrix} = \begin{bmatrix}26.693\\ 1.113\\ -0.127\end{bmatrix} \tag{2}$$

In the diagnostic checking step, we test adequacy and closeness of fit of the model to the data by sample autocorrelation function of the residuals say

$$e_t = y_t - \hat{y}_t, \quad t = 1, 2, ..., N-2 \tag{3}$$

The sample autocorrelation function of the residuals is shown in Fig. 3.

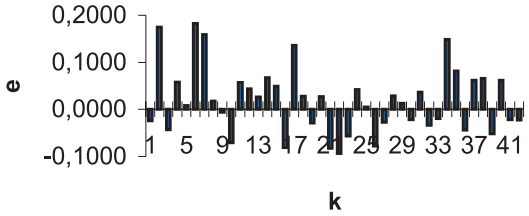

Fig. 3 Sample autocorrelation function of the residuals for model (2)

The modified Box-Pierce statistic Q is used for collectively testing the magnitudes of the residual autocorrelations for insignificance. The statistic is [5]

$$Q = (N - d)\sum_{k=1}^{K} r_{ek}^2 \tag{4}$$

where r_{ek}^2 is the square of the residual autocorrelation coefficients, for lags $k = 1, 2, ..., K$, d is dth differences of the data. For our stock price time series the Box-Pierce statistic for lag $k = 42$ was computed to be 27.78. This value is less than the critical chi square value of 55.7585 (degrees of freedom is $42 - 2 = 40$, $= 0.005$). Hence, we can conclude that the error terms are random and the model (1) is an adequate model.

3. Stock price prediction using adaptive signal processing procedures

In practice, the modelling of a set of data as we shown in Section 2, is a much more complex process than the one of fitting and testing. Most of deterministic methods to signal processing have the goal of representing a given sequence $\{y_t\}$ as the impulse response of a rational linear system [2], [12]. The AR model involves a linear filter with transfer function $H(z)$, where

$$H(z) = [A(z)]^{-1} = \frac{1}{1 + \sum_{k=1}^{p} a_k z^{-k}} \tag{5}$$

and is also known as the all-pole model which has only a nontrivial denominator polynomial, generating the random process $\{y_t\}$ from the white noise $\{\epsilon_t\}$. The linear filter is represented by the inverse of the polynomial $[A(z)]^{-1}$. The difference equation for the input-output relationship for this filter is

$$y_t = \sum_{k=1}^{p} a_k y_{t-k} = \epsilon_t \tag{6}$$

or

$$y_t = \sum_{k=1}^{p} a_k y_{t-k} + \epsilon_t \tag{7}$$

where a_k are the filter parameters that determine the location of the poles of linear filter $H(z)$. To forecast the observation we can

take expectation at origin $t-1$ of the model (7) written at time $t+1$, namely

$$E[y_t] = \hat{y}_t = -\sum_{k=1}^{p} a_k y_{t-k} \tag{8}$$

For selecting the model order p, we will now monitor prediction Mean Square Error (MSE) of the model (8). If the data is truly described by a finite-order AR model, then the theoretical MSE becomes constant once the model order is reached. A realization of this criterion is shown in Fig. 4.

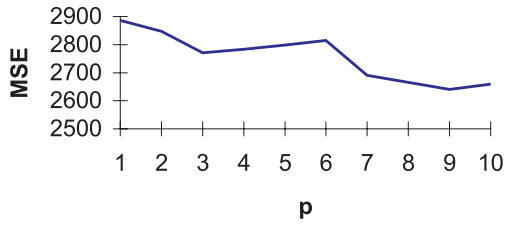

Fig. 4 Graph of MSE (analysis period) versus model order p.

As can be seen from Fig. 4, the MSE's seen to become constant at $p \geq 7$. Then the linear filter (7) has the form

$$y_t = -\sum_{k=1}^{7} a_k y_{t-k} + \epsilon_t \tag{9}$$

The final estimates of predictor parameters (9) are obtained using two of adaptive filtering algorithms in signal processing. The Gradient Lattice (GL) adaptive algorithm and Least Squares Lattice (LSL) form for the parameter estimates of the predictor (9) were used. These algorithms are coded in MATLAB and described in [2]. In this case the process of calculating predictor parameters are updated as each new data point becomes available to track the changing statistics. In Tab. 1 parameters of AR process and the corresponding RMSE's (Root Mean Square Errors) for models (1) and (9) are given.

The RMSE's are called standard deviations of the single-period-ahead forecast errors. For this measure, the AR models estimated by OLS and GL procedures not exceed 5 % limit of the variation coefficient, while the variation coefficient (V) computed as $V = RMSE/\bar{y}$, where $\bar{y}$ is the mean of the stock price time series, for AR model estimated by LSL procedure is 5.59%. Fig. 1 shows the GL prediction results and actual values for stock price time series in both analysis and ex post forecast period. The GL approximations in both intervals visually match the actual stock series quite well. However the corresponding RMSE value for analysis period is 214.48. This is greater than that for ex post-forecast period. The RMSE statistic is here not adequate, because most errors are fairly small, i.e. the model is a good fit to the historical data but there are only first three large errors, these are magnified by using RMSE (since all the errors are squared). This phenomena is produced by GL algorithm itself and this does not influence the ex post forecast errors.

4. Neural network approach

The structure of an ANN is defined by its architecture, its activation functions and learning algorithm. While many variations are possible we suggested an alternative of the most common form of ANN which was suggested and discussed in [8]. This alternative of ANN is pictured in Fig. 5.

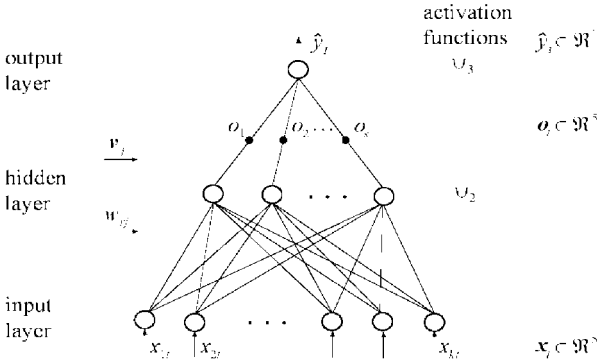

Fig. 5 Fully connected single hidden layer network

Fig. 5 shows a fully connected and strictly hierarchical ANN with variational number of inputs, further variational number of hidden layer units and one output unit. Processing units of the hidden layer have an activation function S - shaped tanh, which produces values of outputs $o_j, j = 1, 2, ..., s$ ranging from -1 to 1. Processing units of hidden layer have the associated weights $w_{rj}, r = 1, 2, ..., k, j = 1, 2, ..., s$. Input data x_r of the ANN are standardized variables. The standardized version of the variables is created in a data-preprocessing unit. The system in a preprocessing unit subtracts the mean of the variable from each observation in the variable and divides the result by the standard deviation of that

OLS, GL and LSL estimates of AR models Tab. 1

Model	Order	Est.proc.	$\hat{\xi}$	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{a}_5$	$\hat{a}_6$	$\hat{a}_7$	RMSE*
(1)	2	OLS	26.639	1.113	−0.127						67.787
(1)	7	OLS	45.930	1.085	0.0861	−0.2531	0.0836	−0.0057	0.2081	−0.2281	76.548
(9)	7	GL		−0.7513	−0.1701	−0.0230	−0.0128	−0.0028	−0.0472	0.0084	68.540
(9)	7	LSL		−0.8941	−0.6672	0.7346	−0.2383	0.1805	−0.5692	0.4470	94.570

*ex post forecast period

variable. After standardization all input variables x_r have values ranging from -1 to 1 and the bias equals to zero. Hidden layer weights w_{rj} are estimated from data according to the learning technique and choice of measure of accuracy in any ANN application. The processing units of hidden layer produce output values o_j such as

$$o_j = \tan H\left(\sum_{r=1}^{s} w_{jr} x_r\right) \quad j = 1, 2, ..., s$$

A dependent variable $\hat{y}$ is produced in an output unit. The output layer unit produces a dependent variable $\hat{y}$ so that the hidden layer outputs $o_j, j = 1, 2, ..., s$ are each multiplied by an additional parameter (weight) estimated from the data. These weights have a clear interpretation. They show how hidden layer outputs o_j contribute to the dependent variable $\hat{y}$ (total model). The dependent variable is transformed to the origin scale in postprocessing unit.

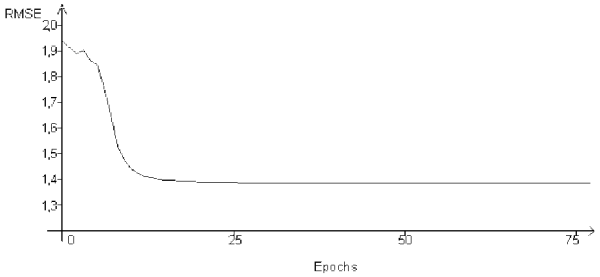

Fig. 6 RMSE's - validation set (normalized data)

Our ANN was trained on the training data set by Back-Propagation algorithm. Periodically, during the training period, the RMSE of the ANN were measured not only on the training set but also on the validation set. The final ANN chosen for the stock price prediction is the one with the lowest error on the validation set (see Fig. 6).

The RMSE's of our predictor models are shown in Tab. 2.

Tab.2

Model	RMSE*
AR(2) - OLS estimates	67.7
AR(7) - OLS estimates	76.5
AR(7) - GL estimates	68.5
AR(7) - LSL estimates	94.6
Neural network	67.2

* Validation set

According to the results of our experiments, the predictor based on the ANN forecasting model is the best, but only slightly better than the AR(2) model. As is stressed in [10], neural networks can outperform standard forecasting procedures at least for certain types of situations. Namely, where the relationship between inputs and outputs are highly nonlinear. Because the results were based on chosen stock price time readings, they were difficult to generalize in other situations. Yet, the results certainly provide a rational way for improvement of forecasting ability in nonlinear economic systems.

Acknowledgement

This work was supported by the Slovak grant foundation under the grant No. 1/0499/03, 1/9183/02.

References:

[1] BARNET, W., HE, Z.: *Unsolved econometric problems in nonlinearity, chaos, and bifurcation.* CEJOR (2001) 9, 147-182, Heidelberg: Physica Verlag, 147-182

[2] BAYAHAN, G. M.: *Sales Forecasting Using Adaptive Signal Processing Algorithms.* Neural Network World 4-5/1997, Vol. 7, pp. 579-589

[3] BELL, T. B., RIBAR, G. S., VERCHIO, J. R.: *Neural nets vs. Logistic regression: A comparison of each model's ability to predict commercial bank failures.* Presented at the 1990 Deloitte & Touche Auditing Symposium. University of Kansas.

[4] BOX, G. E. P., JENKINS, G. M.: *Time Series Analysis, Forecasting and Control,* Holden-Day, San Francisco, CA, 1970

[5] BOX, G. E., PIERCE, D. A.: *"Distribution of Residual Autocorrelation in Autoregressive-Integrated Moving Average Time Series Models".* Journal of the American Statistical Association, vol. 64, 1977

[6] FISHMAN, M. D., BARR, LOICK, W.: *Using neural networks in market analysis.* Technical analysis of Stocks and Commodities. 18-25

[7] HILL, T. L., MARQUEZ, M., O'CONNOR, REMUS, W.: *Neural network models for forecasting and decision making.* IntenationalInternational Journal of Forecasting 10, 1994, 5-15

[8] MARČEK, D.: *Neural networks and fuzzy time series with applications in economics.* Silesian University, 2002 (in Czech)

[9] MARČEK, D.: *Stock Price Prediction Using Autoregressive Models and Signal Processing Procedures.* Proceedings of the 16th Conference MME'98, Cheb 8.-10.9.1998, 114-121

[10] PELIKÁN, E.: *Time series forecasting by neural networks.* International Conference „Artificial neural networks and their application possibilities". TU in Košice, November 18.-19. 1996, 96-101

[11] SURKAN, A., SINGLETON, J.: *Neural networks for bond rating improved by multiple hidden layers.* International Joint Conference on Neural Networks. Vol 2, San Diego, CA, IEEE, New York, NY, 157-162

[12] THERIEN, C. W.: *Discrete Random Signals and Statistical Signal Processing.* Prentice-Hall, USA, 1992.