Anna Górecka
Maciej Szmit

Exchange Rates Prediction by Arima and Neural Networks Models

Article from Artificial Neural Networks Workshop on 47th International Atlantic Economic Conference

The newest history of Warsaw Stock Exchange begun in March 1991 when Act on Public Trading in Securities and Trust Fund was adopted. The State Treasury established the Warsaw Stock Exchange (WSE in polish WGPW) joint-stock company in April 1991. At the same time, the Polish Securities Commission, with its chairman appointed by the Prime Minister, was created. Both the structure and the legal regulations of the Polish capital market were patterned after the most modern and efficient system based on French experience and was adopted and implemented with the help of experts from Societe des Bourses Francaises and the French Depository (SICOVAM).

At the first period of its history of WSE was using the ? call market? system. It is also known as the ? Single-price auction? (French: par casier or German: Einheitskurs). Its main feature is that a single price per security emerges at each session as a result of the orders submitted. Transaction can be concluded even for one individual security. In their orders, clients of brokerage houses define the quantity and the price of securities. In his order investor can indicate the price limit or "at the market". The validity of an order can not exceed the end of the next month. The price on the equities market can only be higher (upper limitation) or lower (lower limitation) from the previous session's price by a maximum of 10%. For bonds maximum price change is 5 percentage points.

The main index of WSE is named Warsaw Stock Index (WSI in polish WIG) is Total return index, which includes dividends and pre-emptive rights (subscription rights). Base value of it was 1000 and the highest value in history was about 20000. Construction of WIG based on the following formula:

WARSAW STOCK INDEX (WIG) FORMULA

$$WIG(t) = \frac{M(t)}{M(0)} \bullet K(t) \bullet 1000$$

where:

M(t) – capitalization of index portfolio on session "*t*"
M(0) – capitalization of index portfolio on the base date (16 April 1991)
K(t) – chain index factor for session "*t*"

In our research we used the multi-layer ANN trained by backpropagation algorithm to predict exchange rates on Warsaw Stock Exchange. We compared forecasts obtained by ANN to the ones generated by for Auto-Regressive Integrated Moving Average Process - ARIMA.

The exchange rates prediction has been very interesting not only for econometricians, nor statisticians. There are a lot of techniques, which are used to predict the rates. We can distinguish four groups of them:

- Technical analysis
- Fundamental analysis
- Econometric analysis
- Artificial intelligence methods

The methods belong to two last groups: ARIMA is classical statistical (Econometric in this case) model and ANN are is AI of course. The acronym ARIMA stands for "Auto-Regressive Integrated Moving Average". Classical ARIMA is a linear model based on least-squares estimation of AR coefficients and it can be exactly calculated from autocorrelations in a single iteration. A nonseasonal ARIMA model is classified as an "ARIMA(p,d,q)" model, where:

- p is the number of autoregressive lags,
- d is the number of differences, and
- q is the number of lagged forecast errors in the prediction equation.

ARIMA

$$\varphi\,(B)\nabla^d z_t = \Theta(B)a_t$$

where:

B - backward operator

$$Bz_t = z_{t-1}$$

φ - p-period autoregressive operator

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p,$$

so

$$\varphi(B)z_t = z_t - \varphi_1 z_{t-1} - \varphi_2 z_{t-2} - \ldots - \varphi_p z_{t-p}$$

∇^d - d-order differentation operator

$$\nabla z_t = z_t - z_{t-1}$$

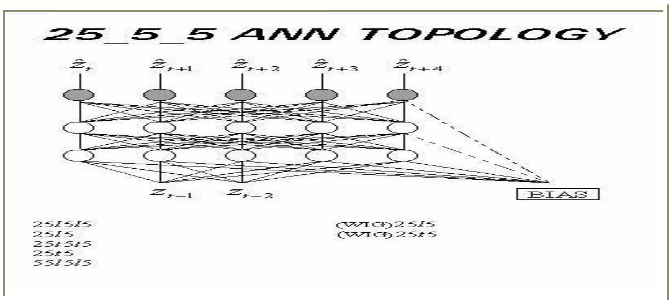
Θ - q-order moving average operator

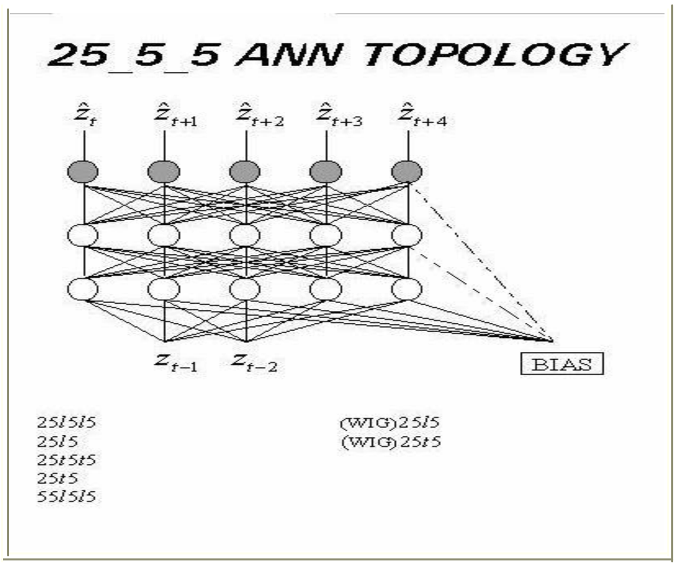
$$\Theta(B) = 1 - \Theta_1 B - \Theta_2 B^2 - \ldots - \Theta_q B^q,$$

or

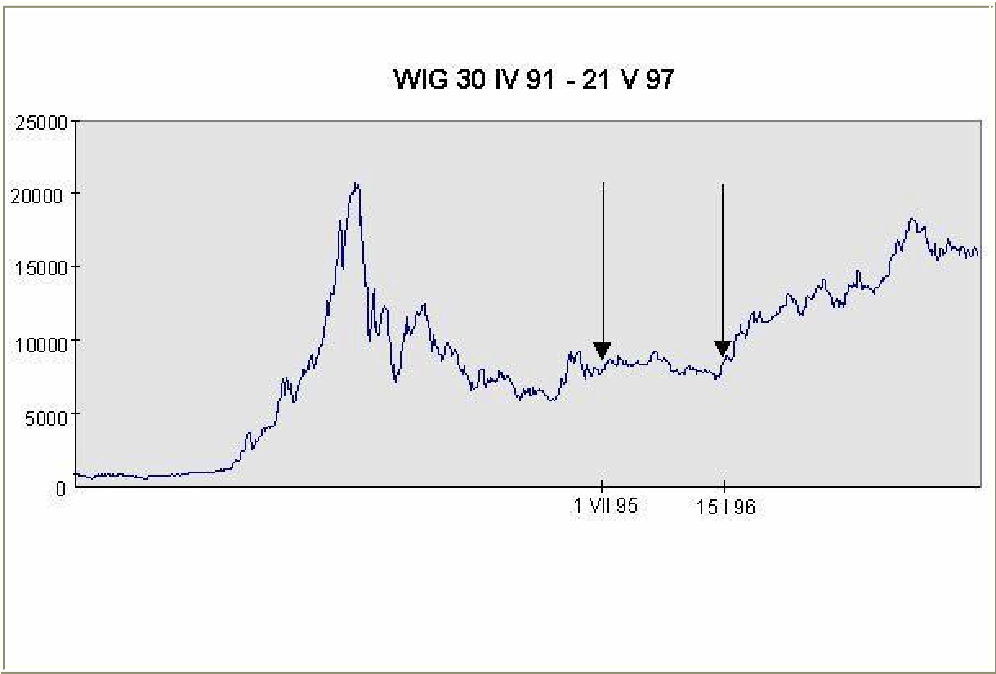
$$\Theta(B)a_t = a_t - \Theta_1 a_{t-1} - \Theta_2 a_{t-2} - \ldots - \Theta_q a_{t-q}.$$

ARIMA is a very effective model and it is not too complex to use it in a lot of applications. More information about ARIMA model we can find in [Box Jenkins].
We also used a few ANN with multi-layers structure: with different number of layers, with different type of activation functions, with different number of neurons in layers and with different input information. In the first experiment we used five kinds of ANNs named like follow:



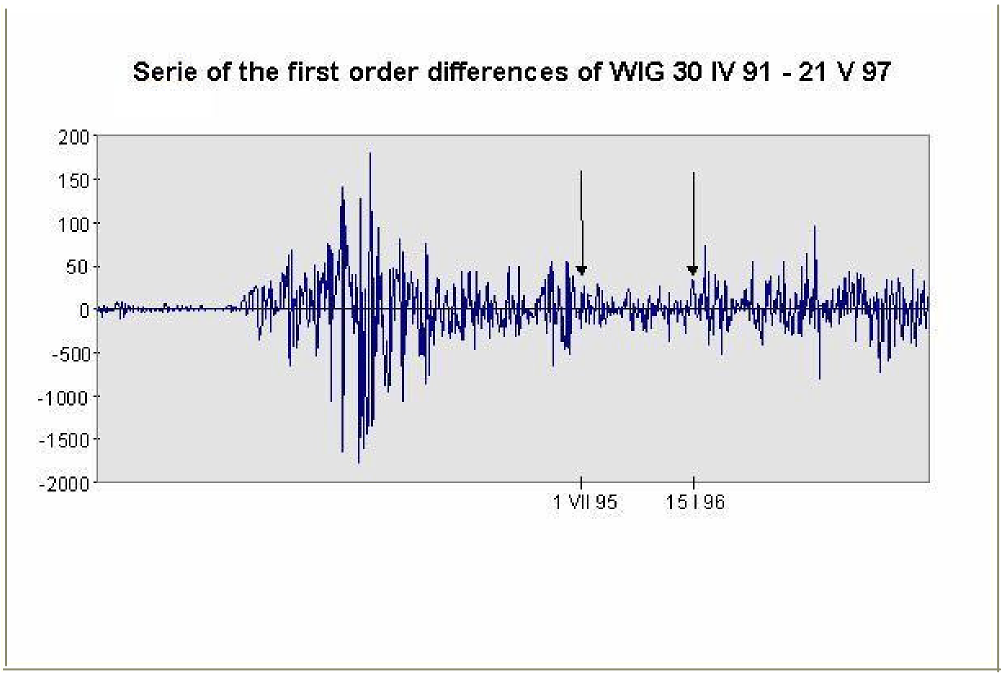


for example 25l5l5 means: two linear neurons in the input layer, five neurons with logistic activation function in the first hidden layer, five neurons with logistic activation function in the second hidden layer and five linear neurons in the output layer - etc. (t means hiperbolic tangent),
The input values were actual lagged by 1 and 2 (up to 5) WSE sessions exchange rates of modeled security. In all the neural networks we had 5 output variables that represent exchange rates on zero, one, two up to four days after.
In the second experiment we add an extraordinary information to input (only for 25t5 and 25l5 ANN). It was a one day before value of the index of Warsaw Stock Exhange (Warsaw Stock Index). In our experiments we tried to predict values of five random securities. The examined period (1.07.95-15.01.96) was characterized by stable behavior of rates



The parameter of a differentiation in the ARIMA model was founded experimentally as follow

WIG	EXBUD	KABLE	KROSNO	TONSIL	PRÓCHNIK
(1,1,0)	(2,1,2)	(1,1,1)	(1,1,2)	(2,1,2)	(2,1,2)



For tested ANN in the first experiment we use input information the same as for ARIMA. Mathematically the difference between the two models (ARIMA and ANN) is that the first is based on linear autoregression and the second ? on non linear. To prevent for errors of ANNs forecasts we repeated learning process (and prediction) three times for every kind of network. Sample results for ANN 25t5t5 are as follow:

The 25t5t5 ANN prognoses

			Value of rate					Prognose					APE				
	Company	Iter	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1	Exbud	294	34,5	34,6	34	33	33,5	34,1	36,3	38,8	40,7	42	1%	5%	14%	23%	25%
2	Exbud	776	34,5	34,6	34	33	33,5	36,3	38,1	39,7	40,5	40,9	5%	10%	17%	23%	22%
3	Exbud	202	34,5	34,6	34	33	33,5	30,9	31,7	32,2	32,6	32,6	10%	8%	5%	1%	3%
																MAPE:	**12%**
1	Kable	713	34,6	35	34,6	34,2	35,4	35,1	34,9	34,7	34,5	34,2	1%	0%	0%	1%	3%
2	Kable	234	34,6	35	34,6	34,2	35,4	35,7	35,5	35,4	35,1	34,8	3%	1%	2%	3%	2%
3	Kable	298	34,6	35	34,6	34,2	35,4	35,6	35,5	35,4	35,5	35,3	3%	1%	2%	4%	0%
																MAPE:	**2%**
1	Krosno	475	44	44,9	44	45	44	42,4	42,4	42,5	42,7	42,6	4%	6%	3%	5%	3%
2	Krosno	1355	44	44,9	44	45	44	-257	-315	-360	-377	-412	684%	801%	918%	938%	1036%
3	Krosno	532	44	44,9	44	45	44	43,7	43,3	43,2	43	42,8	1%	3%	2%	5%	3%
																MAPE:	**3%**
1	Prochnik	1099	15,1	14,9	14,3	14,3	14,5	13	12,7	12,5	12,2	12,4	14%	15%	12%	14%	15%
2	Prochnik	771	15,1	14,9	14,3	14,3	14,5	12,9	12,5	12,2	11,9	11,8	15%	16%	15%	16%	19%
3	Prochnik	59	15,1	14,9	14,3	14,3	14,5	14,2	14,2	14,1	14,1	14,1	6%	5%	2%	1%	3%
																MAPE:	**11%**
1	Tonsil	874	18,7	19,4	19,3	15,7	16,5	18	17,5	17,2	16,8	16	4%	10%	11%	7%	3%
2	Tonsil	183	18,7	19,4	19,3	15,7	16,5	19,7	9,32	30,5	27,6	24,2	5%	52%	58%	76%	47%
3	Tonsil	474	18,7	19,4	19,3	15,7	16,5	15,7	15,5	15,3	15,2	15,3	16%	20%	21%	3%	7%
																MAPE:	**23%**
1	WIG	593	8980	8926	8835	8799	8912	8874	8838	8800	8769	8738	1%	1%	0%	0%	2%
2	WIG	1212	8980	8926	8835	8799	8912	8868	8801	8747	8662	8609	1%	1%	1%	2%	3%
3	WIG	414	8980	8926	8835	8799	8912	8797	8712	8624	8572	8521	2%	2%	2%	3%	4%
																MAPE:	**2%**

Exbud, Kable, Krosno, Próchnik and Tonsil are names of joint stock companies and WIG is Warsaw Stock Index acronym in Polish. Please note that ANN can give very unexpected (and nonsense) results (1036% error in case two of KROSNO and negative value of exchange rate. It was after the finish of learning process, which was conducted by means of backpropagation algorithm with the Levenberg-Marquandt criterion. As a matter of fact, there was the worst result in our experiments. So we have got five tables like that (one for every kind of ANN) in the first experiment, two tables like that for ANN from the second experiment, and one table with ARIMA prediction results. And we could try to find a universal criterion for comparing the quality of our models. We decide to use mean absolute percentage error

(MAPE) for all securities and all forecasts (five days) given as follows

MAPE

$$O_{\acute{s}} = \frac{\sum_{j=1}^{n}\sum_{i=1}^{m}\frac{|p_{ij} - w_j|}{w_j}}{n \cdot m}$$

where:

i - indicator of forecast ($i = 1, ..., m$),
j - indicator day of forecast ($j = 1,...,n$),
m - prediction horizon ($m = 5$)
n - number of forecasts ($n = 3$)
p_{ij} - value of forecast i for day j
w_i -value of real exchange rate in day j

And there are our results:

RESULTS

	251515	25t5t5	2515	25t5	551515	(WIG)2515	(WIG)25t5	ARIMA
Exbud	39%	12%	15%	21%	12%	15%	15%	9%
Kable	2%	2%	2%	2%	3%	4%	3%	1%
Krosno	19%	3%	4%	4%	2%	5%	3%	3%
Prochnik	6%	11%	4%	4%	7%	6%	6%	5%
Tonsil	22%	23%	6%	7%	22%	6%	8%	8%
WIG	2%	2%	2%	1%	1%			1%

As we can see the results for the simplest network are even better then the ones for more complex ANN, and they are comparable with ARIMA. The second conclusion is that the network with more input information are not much better than ANN without it. This is one of the most important advantages of ANN ? it is simply in use (it is not too necessary to try a special, very complicated network topology, extra input information etc.) and it can give results comparable with classical (and more complicated) statistical models.

But there are also bad features ? in our opinion one of the most important is unexpected reaction of network, unexpected output values like the negative values of exchange rates.

A lot of scientists who work with ANN are followers of them. We have to agree that ANN can be a useful tool for economist (especially in starting period of research) but in our opinion it rather cannot replace the classical statistical models.