VOL. 10, NO 22, DECEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

RESEARCH ON THE EFFECTIVENESS OF PREDICTION MODELS IN THE SECURITIES MARKET

Siham Abdulmalik Mohammed Almasani, Valery Ivanovich Finaev and Wadeea Ahmed Abdo Qaid
Southern Federal University, Bolshaya Sadovaya Str, Rostov-on-Don, Russia
E-Mail: siham.almasani@mail.ru

ABSTRACT

The article describes the methods and models of prediction in case of the stock market. There are different methods for assessing and predicting in case of stock market, which are widely used in practice in the present. In this paper, we used the model of Moving Average Convergence/Divergence, prediction model based on a neural network, model of autoregressive integrated moving average (ARIMA) for prediction. We identified the conditions of the use the models as indicators in the case of the market, show examples of application models on the data of the closing prices for the exchange rate from dollar USD to ruble in Russia. In this article analyzes the disadvantages of these models and the causes of these deficiencies. The materials characteristics in this article are analyzing the practical orientation of prediction models in the case of the stock market and interpret the results of mistaken prediction.

Keywords: stock market, state, forecast; securities, modeling, moving averages, neural networks, autoregressive.

1. INTRODUCTION

Securities market or the stock market occupies a separate place in the economy and has a significant impact on the economic situation of enterprises and industries. The effect of Securities market in the redistribution of resources between sectors and investment is to accelerate and optimize the structural changes in the economy.

The task of prediction in the case of the securities market is very important for the present, since the prediction in the case of the securities market is associated with investment activity, This means relationship the prediction with investment the money in order to obtain profit. There are different methods for assessing and predicting in the case of the stock market, which are widely used in practice. All famous methods are heuristic, but in different degrees. Research on the effectiveness of these methods is an important task; the accuracy of predicting in the case of the securities market is directly related to the possibility of profit and loss from investment.

2. MODELS AND METHODS

2.1. Model of moving average convergence/divergence (indicator MACD)

Moving Average Convergence/Divergence (MACD) is a technical Indicator for assessment and prediction of price fluctuations in the stock and currency markets [1-3]. Indicator MACD constructed based on the moving averages, the indicator used to evaluate the existing trend and identifying turning points. When calculating the indicator values take into account the difference between the 26 and 12-period of exponential moving average.

For determine the moment of sale or purchase in the chart, also need to construct a signal line, corresponding 9-period of moving average.

MACD indicator is used in the moments of market fluctuations in a trading range with high amplitude, Moreover, there are rules for the implementation of transactions depending on changes in the value of the MACD indicator and its relationship to the signal line [1,4,5], indicate the weakness of the current uptrend or downtrend [2, 3, 6], and so on.

The calculation of the MACD indicator is carried out by the formulas [3]

$$MACD = EMA\ (CL,\ 12) - EMA\ (CL,\ 26), \qquad (1)$$

$$SIG = EMA\ (MACD,\ 9), \qquad (2)$$

$$Histogram = MACD - SIG, \qquad (3)$$

$$EMA = Price\ (t) * k + EMA(y) * (1-k), \qquad (4)$$

$$k = 2/\ (n+1). \qquad (5)$$

Where

EMA - is the Exponential Moving Average; SIG - the Signal Line of the Indicator; t - The price at the current time; y- The price of the previous time; n - is the number of days in the EMA;

We calculate the MACD indicator for the US dollar (USD) to ruble (RUB). There are daily closing prices for the exchange rate of the US dollar (USD) to ruble (RUB) in Russia from 05.01.2014 to 06.17.2014. Figure-1 shows the result of calculation of the indicator MACD. Where X-axis is specified date range in 34 days,

VOL. 10, NO 22, DECEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

Y-axis indicates the values of the indicator MACD, signal line and histograms.

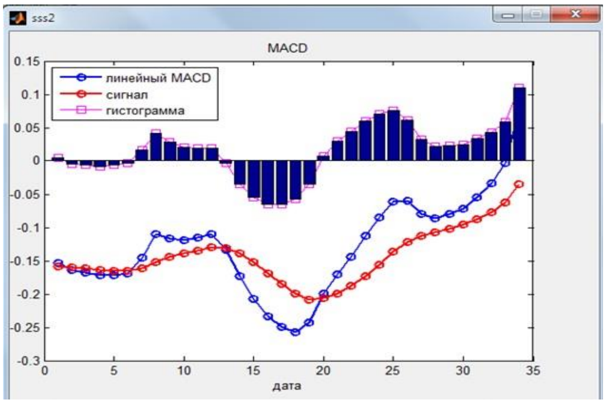

Figure-1. Forecast of securities.

In Figure-1 the line of MACD is above the zero line, so the market has an upward trend. The intersection of the MACD line with the zero line from the bottom to up gives an indication of the purchase of securities. In Table-1 shows the actual closing price. A comparison with the real closing prices indicates that the MACD gives a false signal for purchasing securities.

Table-1. Changes in closing prices in trading.

Date (2014 г.)	Closing price
18.07	35,175
20.07	35,141
21.07	35,175
22.07	35,000
23.07	34,888
24.07	35,087
25.07	35,140
27.07	35,132

The following conclusions: MACD indicator gives good results in stable markets. MACD indicator stops to give adequate instructions and reflect the behavior of price if there were in the markets significant changes.

Disadvantages of indicator MACD:

- The MACD and MACD-Histogram give many false signals;

- The MACD indicator sometimes delays very much in the formation of direction signals.

- The MACD has settings that can be changed to give almost limitless numbers of variations which means results will always differ from person to person. That means the lack of optimal settings. If the scale is large, MACD miss important signals, if the scale is small then there are a lot of false signals.

21.2. The prediction model based on the neural network

Artificial neural networks are mathematical models and their implementation based on the principle of the organization and functioning of biological neural networks. Neural networks are capable of adaptive learning by reactions to positive and negative impacts. The use of models with neural networks in forecasting gives advantages over other methods. Neural networks gives possibility to analyze large volumes interconnected of data, also the volumes of products dependent on a large number of factors. For analyze possible to take into account seasonal variations in the sale of goods with the use of neural networks [7].

Backpropagation Neural Network - is one of the multilayer neural network models. The model name is associated with the learning algorithm. The algorithm involves two passes through all layers of the network: feed-forward and feed-back. First the feed-forward the input vector is passed to feed the input layer of the neural network and then distributed over the network from layer to layer. As a result, a set of output signals is generated, which is a reaction active of network on the input vector. During the feed-back all synaptic weights are passed and adjusted in accordance with rule error correction: the actual output of the network subtracted from desired, as a result of this formed the error signal. Then, this signal is distributed over the network in the opposite direction to the direction of synaptic connections. Figure-2 shows general view network backpropagation.

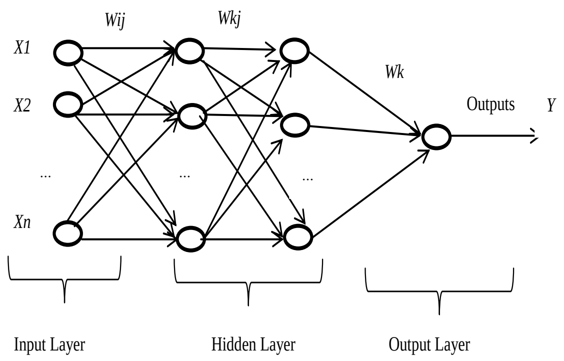

Figure-2. The Form of the backpropagation network.

The output signal of the neural network is defined by the function:

$$Y_t = W_0 + \sum_{j=1}^{n} W_j F\left(\sum_{i=1}^{m} W_{ij} Y_{t-1} + W_{0j}\right) + \delta_t \quad (6)$$

Where: m- is the number of neurons in the input layer, n- is the number of hidden nodes.

VOL. 10, NO 22, DECEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Neural network is training by gradient descent method [8-10]. In each iteration, weights of input and hidden layers change by the formula:

$$Wnew=Wold+\eta\times ERROR\times X, \tag{7}$$

$$ERROR=T-Y, \tag{8}$$

Where: X – input data, W - Weight and η - Learning rate. The type of activation function used sigmoid:

$$f(x)=\frac{1}{1+e^{-x}} \,. \tag{9}$$

The most common formula to calculate the errors in neural networks is a formula of mean square error [8]:

$$MSE=\frac{1}{N}\sum_{i=1}^{N}e_i^{\,2} \,, \tag{10}$$

Where: ei=yi-ti - the difference between the actual and desired output signals at the i-th neuron.

Example: Figure-3 shows the change in the price of the US dollar (USD) to the Russian ruble (RUB) from the first of 2002 to 2014

Figure-3. Prices Dollar USD to Ruble (RU).

For training the neural network was used static data in the period from 01.10.2012 to 30.06.2014 year. The structure of the neural network in (Figure-4) by using the method of iterations, the parameters were selected: the input data: the opening price, low, high, the closing price; output- closing price; percentage of training data (70%); percentage of test data - 30%; learning rate η (0,001); the number of hidden layers - 1; the number of neurons in the hidden layer – 100 and the maximum number of training cycles in 1000.

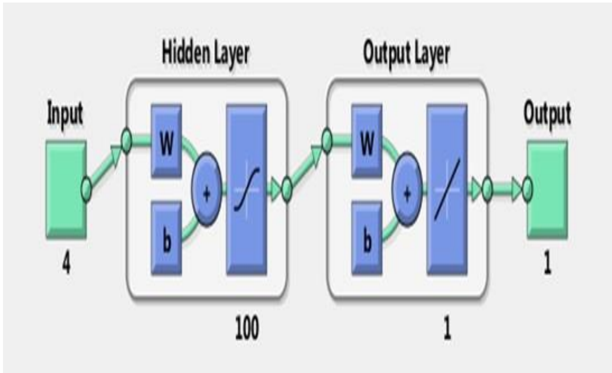

Figure-4. The structure of the neural network.

Data were divided into two groups: for training and testing. This allows us to estimate the accuracy of the neural network work. The training continued for 1000 cycles. Figure-5 shows the results of training and testing in the network. Figure-6 shows the variation of the mean square error for learning.

For check the accuracy of the model was performed forecast for closing price in exchange of USD / RUB from 01.07.2014 to 12.02.2014. Figure 7 shows the result of the forecast. The value of mean square error was 50.94.

Location agent was determined in 20 iterations and it allows us to assess the speed of training the network. The results test for set of examples used only to obtain additional verification of network performance shown in Figure-8.

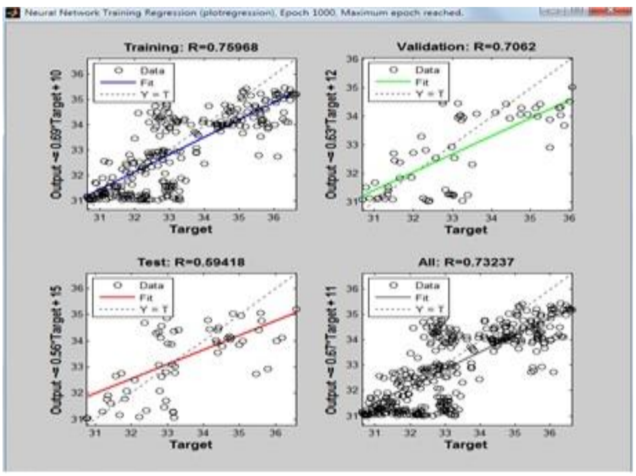

Figure-5. The results of training and testing of the network.

VOL. 10, NO 22, DECEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

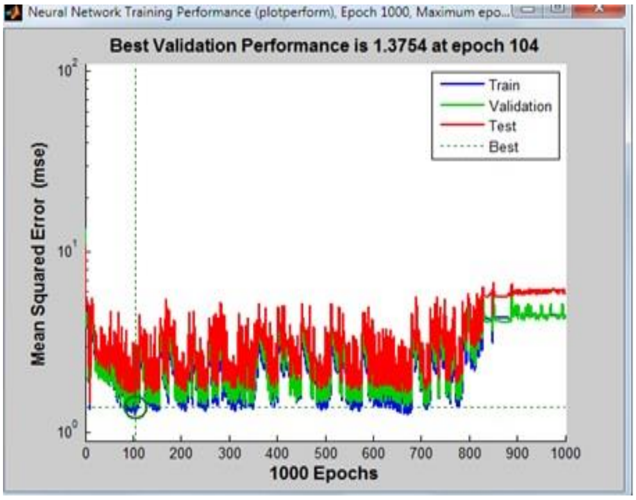

Figure-6. The mean square error of the neural network.

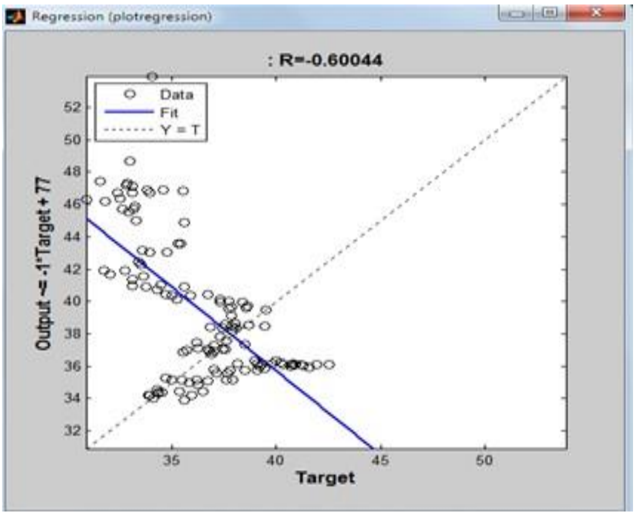

Figure-7. The result of forecast.

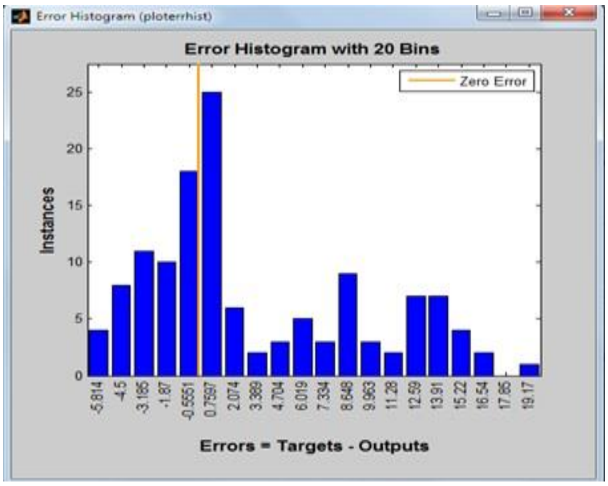

Figure-8. The diagram of neural network deviations from the desired path.

Noted in the previous example in the market occurred significant changes. The accuracy of forecast is not high enough in this experiment to make appropriate decisions in real trading.

The disadvantages of backpropagation network if the price in the market is changing rapidly very high the network can learn incorrectly. There are examples [10-12] the successful application of Backpropagation, but this method is not always effective. The learning process of the neural network can be very long. In the previous example, there were significant changes in the market and the neural network was trained for a long time, but the accuracy of prediction was not high enough. In the complex problems the training of network can take days or weeks, in the end the network may not be trained.

2.3. The autoregressive integrated moving average model (ARIMA)

The autoregressive Integrated Moving Average model can be approximated by both stationary and non-stationary time series [13]. In the ARIMA model for forecasting uses the information of the original series. For example, the ARIMA model for monthly sales volumes identifies temporal structure in the current sales data, which is then used to forecast sales volumes for the next months.

The Box-Jenkins method [14] is the construction of ARIMA models and forecast based on them. ARIMA model specified by set of parameters (p, d, q), where p is the number of autoregressive terms; d is the number of non-seasonal differences needed for stationary; q is the number of lagged forecast errors in the prediction equation. ARIMA model is using the structure of autocorrelation of the data. Consider the features of the algorithm for constructing the model ARIMA (p, d, q) [15]. The construction of ARIMA-model for the study of time series consist three stages: identification of test models; Estimation the parameters of the model and Diagnostic of the adequacy the model; application the model to forecast.

Firstly, necessary get a stationary time series. At this stage, it is recommended to analyze the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the series data. A simple test for stationary is a rapid decay of the ACF values. Also at this stage, can use the statistical tests to find a unit root (augmented Dickey-Fuller or ADF-test) [16].

Accordance for the Dickey-Fuller statistics or estimates ACF if series is non-stationary, then to convert the non-stationary series to a stationary series used operator of taking the successive difference from time series. In the result, determined the value of the parameter d by the number of non-seasonal differences needed for stationary in the model ARIMA (p, d, q).

After get a stationary series investigated, investigated of nature behavior of the sample ACF and

PACF and determined hypothesize about the values of the parameters p (the number of autoregressive terms) and q (the number of lagged forecast errors in the prediction equation). From this information form basic set comprising one, two or more models, depending on the parameters p and q.

After identification the model estimate the model parameters. For this purpose used the maximum likelihood method (MLM) [17]. To check of the adequacy of each model analyzed its series of residual ACF. For adequate the model series the residues similar to white noise, their sample ACF should not differ from zero. To test the hypothesis that the observed data are the realization of "white noise", used Q-statistics [18]. In the result of check may be several models adequate the original data. Therefore, the final selection should take into account two factors: the increase in accuracy (quality of model fitting), the lowest number of model parameters. These requirements are summarized in the information criterion Akaike and Schwarz [19, 20].

With the resulting model can be built accurate and interval prediction for L steps forward. For predicting was chosen recursive mode, initial observation is fixed, and the observation of the test sample is added one by one to the working sample. The forecast horizon is always the same.

To assess the accuracy of the forecast uses standard indicators [21]:

- Mean absolute percentage error (MAPE):

$$MAPE = \frac{100\%}{L}\sum_{t=1}^{L}\frac{x_t - \hat{x}_t}{x_t}. \quad (11)$$

Where Xt - real value, $\hat{X}_t$- predictive value, L - range forecast, Moreover, if the value of MAPE <10%, then the forecast is high accuracy, at 10% <MAPE <20% - the forecast is good, with 20% <MAPE <50% the forecast is satisfactory, while MAPE> 50% - the forecast is bad;

- Signal-to-noise ratio (SER):

$$SER = 10\,ln\left(\frac{\sum_{t=1}^{L} x_t^2}{\sum_{t=1}^{L}(x_t - \hat{x})^2}\right). \quad (12)$$

Example: For the exchange price of the dollar USD to ruble (RUB) (data from 01.10.2012 to 30.06. 2014) the research conducted to identify the ARIMA model with a minimal number of parameters, which allows to make a prediction.

At first checked if the time series is stationary using the Dickey-Fuller statistics. Table-2 shows the test results.

Table-2. The test results for time series by use the Dickey-Fuller statistics.

Test statistics Dickey-Fuller	test statistic	P- value	The critical value	Stati-onarity
No Const	0,7	86,6%	-1,9	False
Const-Only	-1,1	73,5%	-2,9	False
Const + Trend	-2,1	1,6%	-1,6	True
Const+Trend+Trend 2	-1,7	4,3%	-1,6	True

In the Table-2 the p-value is a function of the observed sample results that is used for testing a statistical hypothesis. When the p-value is calculated correctly, such a test is guaranteed to control the Type I error rate to be no greater than α.Testing hypotheses using P-value is an alternative for classical procedure to check through the critical value of the distribution. The level of significance was 5.0%.

According to the values of criterion the ADF-test series is non-stationary, since the null hypothesis of a unit root is confirmed at the 5% level of significance. The ADF-test passed and the difference operation was applied once, the value of the coefficient d = 1, Obtain a model of the form: ARIMA (p, 1, q). Since the series is stationary, make an assessment about the parameters p and q model.

Model ARMA (p, q) represents a model AR (p) and MA (q). For this we use the partial autocorrelation function (PACF) and the autocorrelation function (ACF), respectively. Procedure model AR (p) is selected from PACF and corresponds to the last non-zero coefficient PACF (see Figure-9). The corresponding procedure is used to find the order of the model. Model MA (q) by use only ACF (see Figure-10). Thus, get on the order of the coefficients; get on the model ARIMA (1, 1, 1). Each resulting model is being tested for adequacy through analysis the series of residues of this model. As mentioned

VOL. 10, NO 22, DECEMBER, 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

above (see. Figure-9) that the model adequately describes the process. A series of residues is a random component corresponding to white noise. ACF residues will not be significantly different from zero.

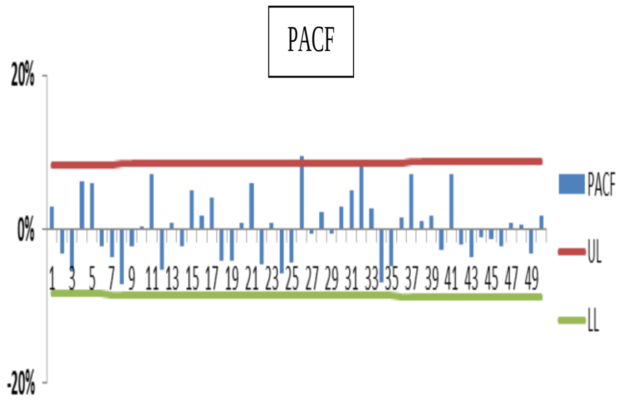

Figure-9. The partial autocorrelation function.

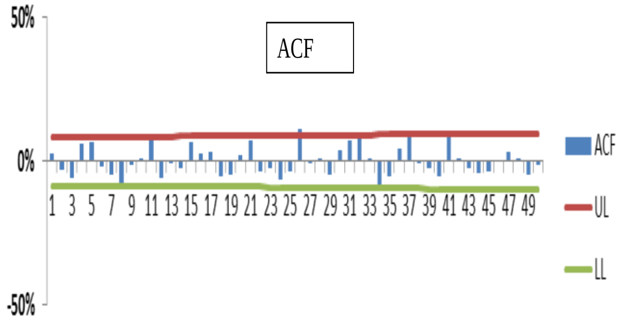

Figure-10. The autocorrelation function

As a result of the calculations was obtained on the variety of models. According to the information criterion SBIC choose the most appropriate model of ARIMA (1, 1, 1) and ARIMA (15, 1, 6), as shown in Table 3 and Table-4.

Table-3. The parameters of model ARIMA (1, 1, 1).

Parameter	***M***	***φ₁***	***θ₁***	***Σ***	***D***
Value	0,01	-0,13	0,17	0,16	1

In Figure-11 shows the results of the forecast for 5 months from 01/07/2014 to 02/12/2014.

Table-4. The Parameters of Model ARIMA (15, 1, 6).

Param	**Value**	**Param**	**Value**	**Param**	**Value**
M	0,01	φ_8	-0,05	θ_1	0,26
φ_1	-0,23	φ_9	-0,02	θ_2	0,35
φ_2	-0,35	φ_{10}	-0,10	θ_3	-0,10
φ_3	0,07	φ_{11}	0,05	θ_4	-0,04
φ_4	0,08	φ_{12}	0,02	θ_5	-0,37
φ_5	0,42	φ_{13}	0,07	θ_6	-0,77
φ_6	0,75	φ_{14}	0,00	σ	0,15
φ_7	-0,04	φ_{15}	0,06	d	1

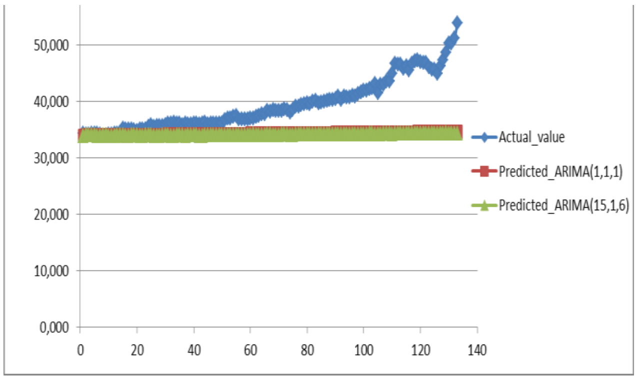

Figure-11. The result of the prediction.

In the Table-5 shows the results of estimates average, characterizing the quality of the forecast evaluation MAPE. These data were obtained from Figure-11 and by MAPE for model ARIMA (1, 1, 1), found that the minimum evaluation for this model corresponds to good forecast and application this model for perform better for forecast.

The general disadvantage of prediction with these models is that they are independent of the calculation methods, which use the prehistory data. If the market conditions (for example, market volatility or correlation between the assets) are changing dramatically, it will consider these changes only after a certain period of time. Up to this time the predictions will be incorrect.

Table-5. The results of estimated average by MAPE for prediction.

	ARIMA (1,1,1)	**ARIMA** (15,1,6)
Data	MAPE	MAPE
All	11,85937	12,29617332
Jul2014	2,025895	2,194185694
Aug2014	5,482338	5,927832108
Sep2014	9,692651	10,22927361
Oct2014	15,76247	16,3081777
Nov2014	26,01704	26,50898117

These models work well in the case of steady markets and no longer adequately reflect the behavior of prices in the markets, when there are significant changes. Indeed, Table 5 shows the result of the quality of the forecast by evaluation MAPE for November 2014 is 26.02% because there are significant changes in the market. If these changes are continuing on in the market, then the prognostic evaluation by MAPE for December 2014 will increase and may be over 50%.

The disadvantages models ARIMA, Requires a relatively large amount of raw data; If the data is periodic, for example, the seasonal period of S = 12, the observation of one full year will be actually one value of the seasonal data, rather than twelve values; In the application of ARIMA model for non-seasonal data need about 40 or more of observations; In the construction of the ARIMA model for seasonal data need observation about 6 to 10 years, depending on the value of the seasonal period.

There is no easy way to adjust the parameters of the models ARIMA, such as in some smoothing methods, when used the new data. The model has to rebuild a periodically completely accounts, and sometimes need to select a completely new model. For building an adequate ARIMA model requires often a lot of time and resource. The cost of construction the ARIMA models, the calculations of run-time and the volumes required of databases can be significantly higher than traditional forecasting methods, such as smoothing.

3. CONCLUSIONS

In this article researched of possibility the model of Moving Average Convergence/Divergence (MACD), the prediction model based on neural network models and autoregressive Integrated Moving Average model to forecast in case of the stock market, for example, the fluctuations in the dollar USD to the ruble in Russia.

Studies have shown the shortcomings in the models for the problem of prediction in unstable market conditions, high rate of change in prices. Thus, this shows the limited conditions for application the prediction models for predicting in the case of the stock market. Therefore, it is necessary to develop methods based on expert knowledge to solve problems of prediction.

REFERENCES

[1] Seyed Hadi Mir Yazdi and Ziba Habibi Lashkari, 2013. Technical analysis of Forex by MACD Indicator. International Journal of Humanities and Management Sciences (IJHMS). 1(2): 159-165.

[2] Terence Tai-Leung Chong and Wing-Kam Ng. 2008. Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30. Applied Economic Letters. 15(14): 1111-1114.

[3] Gerald Appel, E.D. 2008. Understanding MACD (Moving Average Convergence Divergence). Traders Press, Inc.

[4] Yazdi S.H.M. and Z.H. Lashkary. 2012. Technical analysis of FOREX by P-SAP Indicator. International Islamic Accounting and Finance Conference (IIAFC) Malaysia.

[5] Li J. and E.P.K. Tsang. 1999. Improving technical analysis predictions: an application of genetic programming. Proceedings of the 12th International Florida Artificial Intelligence Research Society Conference USA. pp. 108-112.

[6] Neely C., P. Weller and R. Dittmar. 1997. Is technical analysis in the foreign exchange market profitable? A genetic programming approach. Journal of Financial and Quantitative Analysis. 32(04): 405-426.

[7] C.L. Dunis and M. 2002. Williams. Modelling and trading the eur/usd exchange rate: Do neural network models perform better?Derivatives Use. Trading and Regulation. 8(3): 211-239.

[8] L. Peng, Yuehui Chen and A. Abraham. 2006. Exchange rate forecasting using flexible neural trees. Lecture Notes on Computer Science. 3973, pp. 518-523.

[9] P. Tenti. 1996. Forecasting foreign exchange rates using recurrent neural networks. Applied Artificial Intelligence. 10, pp.567-58.

[10] Dase R.K. and Pawar D.D. 2010. Application of Artificial Neural Network for stock market

predictions: A review of literature. International Journal of Machine Intelligence. 2, pp. 14-17.

[11] M. Zelaic. 1998. Neural Network Applications in Stock Market Prediction A Methodology Analysis. Proc. of 9th Intl, Conf. Information and Intelligent Systems.

[12] Yao J and Tan C.L. 2000. A case study on using neural networks to perform technical forecasting of forex. Neurocomputing. 34(1-4): 79-98.

[13] Pai C.Lin. 2005. A hybrid ARIMA and support vector machines model in stock price forecasting. Omega. 33(6): 497-505.

[14] George E. P. Box, Gwilym M. Jenkins and Gregory C. Reinsel. 2008. Time Series Analysis: Forecasting and Control. 4th Edition, ISBN: 978-0-470-27284-8, 784 pp. in Canada.

[15] J. Sterba and Hilovska. 2010. The Implementation of Hybrid ARIMA Neural Network Prediction Model for Aggregate Water Consumption Prediction. Aplimat-Journal of Applied Mathematics. 3(3): 123-131.

[16] C. Javier, E. Rosario, J.N. Francisco and J.C. Antonio. 2003. ARIMA Models to Predict Next Electricity Price. IEEE Transactions on Power Systems. 18(3): 1014-1020.

[17] Ayodele A. Adebiyi, Aderemi O. Adewumi and Charles K. Ayo. 2014. Stock Price Prediction Using the ARIMA Model. UKSim-AMSS 16th International Conference on Computer Modelling and Simulation IEEE. 15(4): 105-111.

[18] B. Uma Devi, D.Sundar and Dr. P. Alli. 2013. An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50. International Journal of Data Mining and Knowledge Management Process (IJDKP). 3(1): 65-78.

[19] Shangkun Deng, Mitsubuchi.T, Shioda. K, Shimada. T and Sakurai. A. 2011. Combining Technical Analysis with Sentiment Analysis for Stock Price Prediction. Dependable, Autonomic and Secure Computing (DASC) IEEE Ninth International Conference on Sydney, NSW. pp. 800-807.

[20] Hannan. E.J. 1980. The estimation of the order of an ARMA process. The Annals of Statistics. 8(5): 1071 - 1081.

[21] Wurtz. D, Y. Chalabi and L. Luksan. 2006. Parameter estimation of ARMA models with GARCH/APARCH errors: An R and SPlus software implementation. Journal of Statistical Software. pp. 1-41.