Forecasting Stock Price using Hybrid Model based on Wavelet Transform in Tehran and New York Stock Market

Vahid Vafaei Ghaeini
Ms. in Financial Engineering, Department of Industrial Engineering and Management Systems,
Amirkabir University, Tehran, Iran
(Corresponding Author)
vahidvafaee7@yahoo.com

Alimohammad Kimiagari
Associate Prof in Industrial Engineering and Management System, Department of Industrial Engineering and Management Systems, Amirkabir University, Tehran, Iran
kimiagar@aut.ac.ir

Mostafa Jafarzadeh Atrabi
Ms. Student in Financial Engineering, Department of Industrial Engineering and Management Systems,
Amirkabir University, Tehran, Iran
Jafarzadeh.m@aut.ac.ir

ABSTRACT

Forecasting financial markets is an important issue in finance area and research studies. On one hand, the importance of prediction, and on the other hand, its complexity, have led to huge number of researches which have proposed many forecasting methods in this area. In this study, we propose a hybrid model including Wavelet Transform, ARMA-GARCH and Artificial Neural Network (ANN) for single-period and multi-period forecasting of stock market price in different markets. At first, we decompose time series into detail and approximate series with wavelet transform, and then we used ARMA-GARCH and ANN models to forecast detail and approximate series, respectively. In addition to the approximate series, we use some technical indexes in this model to improve our ANN model. To evaluate the proposed model in forecasting stock price, we compare our model with ANN, ARIMA-GARCH and ARIMA-ANN models on Tehran and New York Stock Exchange (NYSE) historical prices. The results of study show that the proposed model has better performance in single-period forecasting on Tehran and New York market rather than other models.

Keywords:

ARIMA-GARCH, Forecasting, Neural Network, Wavelet Transform.

1. Introduction

Since people cannot properly understand what is happening in a few moments later, predicting what will happen in the future is an issue in many areas, especially in financial markets.

One of the most significant information for investors in stock market is stock price. It's obvious that if investors be able to predict stock price movements, they can improve their investment returns. Stock prices are basically dynamic, non-linear, nonparametric, and chaotic in nature. For this reason, it's crucial for investors to employ the time series which are non-stationary and noisy.

In other words, there are lots of factors, such as political events, company policies, quarterly earnings reports, general economic conditions, commodity price index, bank rates, investor expectations, and investor psychological factors which affect stock price movements.

On one hand, the importance of prediction, and on the other hand, its complexity, have led to huge number of researches which have proposed many forecasting methods in this area. Among these methods, two methods of Artificial Neural Networks (ANNs) and time series methods are so well-known. Two of the most important time series methods include Autoregressive Integrated Moving Average (ARIMA) model and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model (Bowden & Payne, 2008; Contreras, Espinola, Nogales, & Conejo, 2002; Doulai & Cahill, 2001; Garcia, Contreras, Akkeren, & Garcia, 2005). Major limitation of this models is linearity assumption which make them inappropriate for nonlinear time series modeling. ANNs models which doesn't have linearity assumption, have been shown that are outperform time series model in nonlinear time series. Lots of papers have used neural network to predict stock trends(Baba & Kozaki, 1992; Chenoweth & Obradović, 1996; Fernández-Rodríguez, González-Martel, & Sosvilla-Rivero, 2000; Ghiassi, Saidane, & Zimbra, 2005; Hyup Roh, 2007; Kaastra & Boyd, 1999; Lendasse, Bodt, Wertz, & Verleysen, 2000).

2. Literature Review

Although these models have very well performance there is a theory which has confirmed that there isn't any specific model to have good performance in all conditions. In other words, each model has its own weaknesses alongside its strengths. Using hybrid models can cover these weaknesses and lead to predictions with less error. For this reason, development of hybrid models has been fast in recent years. The use of hybrid models based on artificial neural network and time series models regarding to the strengths and weaknesses of each one, has been widely studied recently. Zhang (2003) proposed a hybrid model to predict time series based on ARIMA and artificial neural network models. His study was a beginning for hybrid models and researches in this field. At first, he used ARIMA to forecast time series, and since this model can't trap non-linear patterns, he considered residual amount left by that as non-linear section of series in artificial neural network model. Finally, he combined the outcomes obtained from both models together. The results of this study showed that the ARIMA-ANN hybrid model had a much better performance rather than either of these two models, lonely(Zhang, 2003). After Zhang, some studies based on his model, have used a combination of statistical and artificial intelligence models to forecast different time series(Adhikari & Agrawal, 2014; Chen & Wang, 2007; Khashei & Bijari, 2011; Pai & Lin, 2005; Yu, Wang, & Lai, 2005). Similarly, Babu and Reddy (2014) based on Zhang's study, divided time series into two parts with high and low fluctuation using moving average model. They obtained low fluctuation part of time series using moving averaging of the original time series, and, high fluctuation part through subtracting low fluctuation part from the original time series. The low fluctuation part and high fluctuation part were predicted by ARIMA and ANN models, respectively. Finally, they combined the outcomes obtained from both models together. They compared their proposed model with ARIMA, ANN, and Zhang's models. The results showed that their hybrid model had better performance rather than other models(Babu & Reddy, 2014).

In order to provide a robust price forecasting method, wavelet transform have been utilized because they can produce a good local representation of the signal in both time and frequency domains. in the recent year, many forecasting model based on wavelet transform have been proposed(Al Wadi, Ismail, Alkhahazaleh, & Karim, 2011; Chang & Fan, 2008; Dai & Lu, 2008; Jammazi & Aloui, 2012; Joo & Kim, 2015; Khandelwal, Adhikari, & Verma, 2015; Tan, Zhang, Wang, & Xu, 2010; Wang, Wang, Zhang, & Guo, 2011). Dai and Lu (2008) used Wavelet transform to predict time series. They presented a new model by combining wavelet transform and SVM for predicting the Nikkei 225 index. Their proposed model included two stages. In first stage, the wavelet transform was used to split the time series into sub series with different scales. In second stage, the sub series obtained from previous stage were provided for SVM model for prediction. They used wavelet transform with function db1 to db6 to split time series. In this way, each time series was split into one to six levels and considered the function and decomposition level with the least error as final answer. The results showed that their model had better performance than random walk and SVM models(Dai & Lu, 2008). Subsequently, Tan et al. (2010) split the historical time series into an approximate series and some detail series by using wavelet transform with db4 function. They predicted the approximate series with ARIMA-GARCH model and detail series with ARIMA model and then combined them(Tan et al., 2010). Similar to Jammazi and Aloui (2012), Joo and Kim (2015) split the time series by wavelet transform from one to six levels to determine a suitable level for decomposition, and then select the best level based on the Minimum Absolute Percentage Error (MAPE) as the optimal level(Jammazi & Aloui, 2012; Joo & Kim, 2015). Khandelwal (2015) et al. introduced a hybrid method for predicting time series based on wavelet transform. In the first step, they split time series using wavelet transform into an approximate part and some detail parts. In the next step, they predicted the approximate time series by artificial neural network and ARIMA model for detail series.

At the last step, the results of these two models were combined. Their model test on Zhang's paper time series showed that their hybrid model had better prediction accuracy than the hybrid model of Zhang, neural network and ARIMA(Khandelwal et al., 2015).

In this paper, a novel hybrid forecasting model is proposed. At first, we use wavelet transform to decompose time series into detail and approximate series, and then we used ARMA-GARCH and ANN models to forecast detail and approximate series, respectively. In addition to the approximate series, we use some technical indexes to improve our ANN model. In order to recognize and determine the optimum wavelet function and level decomposition, we run our model with four wavelet functions in level two to five. To the best of our knowledge, this method has never been presented in the literature.

The remainder of the paper is organized as follows: Section 2 reviewed previous studies. Section 3 outlines the proposed method. Section 4 presents results of the study and Section 5 provides some relevant conclusions.

3. Methodology

Before introducing the proposed model, we present a brief overview of Wavelet transform, ARIMA, GARCH, and artificial neural networks models.

3.1. Wavelet transform

The most important concept in wavelet analysis begins with the selection of a suitable wavelet function (mother wavelet). The wavelet can be defined as function ψ (t) with mean zero as follows [23]:

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{1}$$

A series can be divided into many series by transmission and displacement parameters a and b, respectively.

$$\psi_{a.b}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}) \tag{2}$$

Therefore, the converted series f (t) is expressed by wavelet transform with the scale a and displacement b as follows:

$$wf(b.a) = \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(t)\psi(\frac{t-b}{a})dt \tag{3}$$

The original series f (t) can also be rebuilt by inverse wavelet transform.

$$f(t) = \int_{0}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{a^2} wf(b.a)\psi_{a.b}(t)dbda \tag{4}$$

3.2. ARIMA

The ARIMA model which introduced by Box and Jenkins, is one of the most important and effective statistical models for predicting financial time series. This model states that future values of a time series represent a linear combination of past values and residuals. An ARMA model (p, d, q) is expressed as follows:

$$\emptyset B(1-B)^{d} y_t = \theta B \varepsilon_t \tag{5}$$

Which y_t is the actual observations, ε_t is the white noise, d is the order of differentiation, B is a backward shift operator which is defined as $BX_t = X_{t-1}$, φB is the self-autoregressive operator which is defined as $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p$, and $\theta(B)$ is the moving average operator which is defined as $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_p B^p$.

To implement the ARMA model, at first, we examine the time series condition and, in non-stationary case, make it stationary by differentiating and determining the order of d. Then, using the ACF and PACF tests, we set the values of p and q. In the next step, the values of the parameters are estimated and the significance of each of them is investigated. In case of insignificance, the parameters are deleted from the model and other parameters are re-estimated. For further details, see the reference [24].

3.3. GARCH

GARCH models are used to predict conditional variance values. Models of this family are more complex than the time series models. Models in this class do not use standard deviations, but use the maximum likelihood to formulate conditional variance. These models are obtained by formulating the conditional variance of past fluctuations, and future amount of this parameter can predict easily. The first model of this category was introduced by the Angel in 1982 which by ARCH(q) model set conditional variance as a function of the squared values of returns in the previous q periods.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_t^2 \tag{6}$$

$$\varepsilon_t = \sigma_t z_t \tag{7}$$

To improve this model, Bollerslev and Taylor introduced the GARCH. This model also considered the past values (p) of the conditional variance in the model.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_t^2 + \sum_{i=1}^{p} \beta_i \sigma_t^2 \tag{8}$$

In applications, it has been shown that GARCH model is better than ARCH model and is more suitable for many financial time series.

Improved models in the following years tried to overcome the deficiencies in the GARCH model. EGARCH (Exponential GARCH) was the next model that formulated the conditional variance in logarithmic state and distinguished between positive and negative shocks in determining future fluctuations. In addition to EGARCH, TGARCH (Threshold-GARCH), GJR-GARCH, and QGAECH models and some other models have been introduced to consider the non-symmetry feature of positive and negative shocks impact on data volatility [24].

3.4. Artificial Neural Network

Neural network is one of the models that has been highly regarded in recent years due to its ability to trap nonlinear patterns. Neural Networks are a combination of simple operating elements that operate in parallel. These elements are inspired by the human neural network. The processing elements in the neural network are neurons, nodes and layers [25].

As the network function is determined by complex relations in nature, we also can train the neural network in order to improve a specific function by assigning a certain amount to the relations (weights) between the elements (neurons). Usually, a forward-looking artificial neural network with a hidden layer is used to predict future values. The output of the neural network p in q can be defined as follows:

$$y_t = \emptyset_0 + \sum_{j=1}^{q} \emptyset_j \, g\left(\theta_{0j} + \sum_{i=1}^{p} \theta_{ij} \, y_{t-i}\right) + \varepsilon_t \tag{9}$$

Which

φ_j (j=0,1,...,q) and θ_{ij} (i=0,1,...,p;j=0,1,...,q) are weights of network, ε_t is residuals amounts, φ_0 and θ_{0j} are biases of networks, and g is transportation function.

3.5. Artificial Neural Network

Due to the nature of the financial time series fluctuations and the linear and nonlinear patterns in them, as well as the combination of models to consider these patterns and reduce their predictive errors, in this study, in the first step, using wavelet transform, the time series of prices is decomposed, and then they will be reconstructed into an approximate (non-linear) series and some details (linear) series (as shown in Figure 1).

In this model, D_1, D_2, ..., D_n are the details time series and A_n is approximate time series. We use ARMA-GARCH time series model to predict details time series and use the three-layer artificial neural network with 10 neurons in the hidden layer to predict the approximate time series. In addition to the approximate series, a set of technical indicators will also be given as input to the neural network, as shown in the table below.

To determine the P and Q values in ARMA model, ACF and PACF tests are used. After estimating the parameters of the ARMA model and checking their significance, again, the ACF and PACF tests are used to investigate the effect of GARCH on the squares of ARMA residuals. If the GARCH effect is accepted, a GARCH (1.1) model is used to predict the residuals of ARMA model. See Figure 2 and 3 for the algorithm and structure of the proposed model.

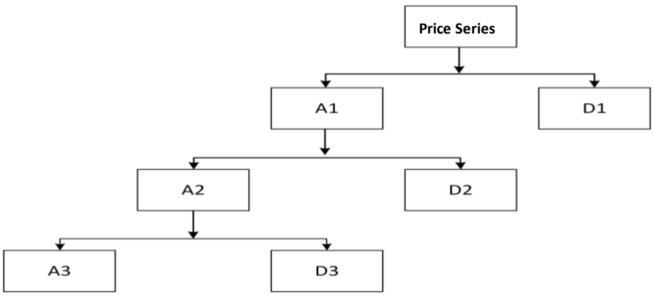

Fig. 1. The decomposed tree with wavelet transform up to 3 levels

Table 1. Input technical variables to neural network along with the approximate time series

Variable	Explanation
Open Price	The first price that trades on that stock each day.
Close Price	The last price of in the day that stock is traded.
High Price	The highest price that stock is traded on that day.
Low Price	The lowest price that stock is traded on that day.
Volume	The amount of stocks traded on the day.
6-day and 10-day Moving Average (EMA6, EMA10)	A simple moving average is used to recognize the trend of stock movements and also to smooth the stock price against stock fluctuations, which can lead to misinterpretations.
12-day and 26-day Exponential Moving Average (EMA12, EMA26)	An exponential moving average, like the simple moving average, is used to recognize trend and price smoothing, except that the closer we get to the current data, their weights will be more.
9-day and 14-day Relative Strength Index (RSI9, RSI14)	The relative strength index compares the importance of recent earnings to losses. This is done to determine the sales saturation area.
14-day Stochastic Line (K,D)	The stochastic line is also used to determine the sales saturation area and indifference saturation area.
9-day Moving Average Convergence Divergence (MACD9)	The moving average of convergence and divergence represents the difference between a fast exponential moving average (short period) and a slow moving average (long period).

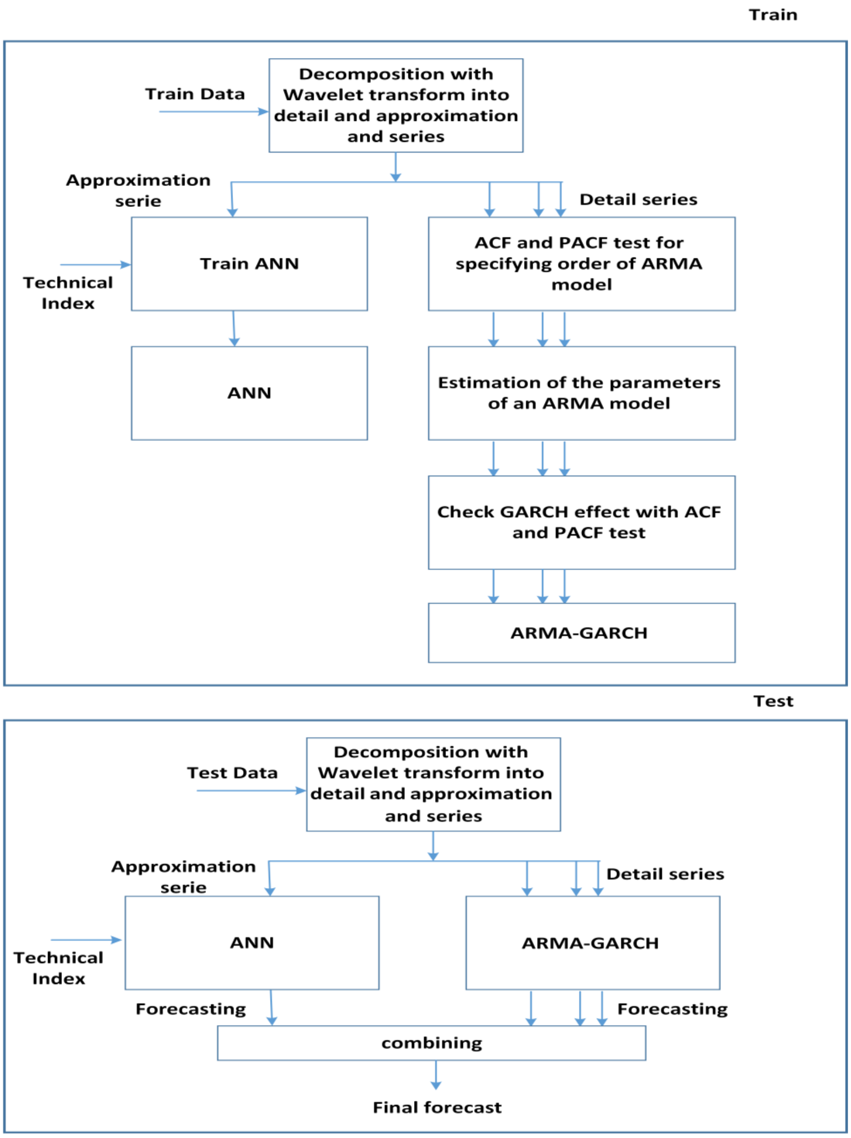

Fig. 2. The proposed model algorithm

Which in the above formulas n represents the number of observations, x_i represents the prediction value in day i, and X_i is the actual value of the price in the i-th day.

In order to investigate and determine the best level of wavelet decomposition and transformation, the proposed model is implemented in 16 modes based on 4 wavelet transformation functions, db2, db4 and haar in four levels of level two, level three, level four, and level five.

The proposed model was used to predict two 1-day and 5-day periods in the Tehran and New York Stock Market and was compared based on MAPE benchmark with three models of neural network, ARIMA-ANN and ARIMA-GARCH. In Table 2 and 5, the results of model implementation in Tehran and New York Stock Exchange are shown for one-day and five-day forecast.

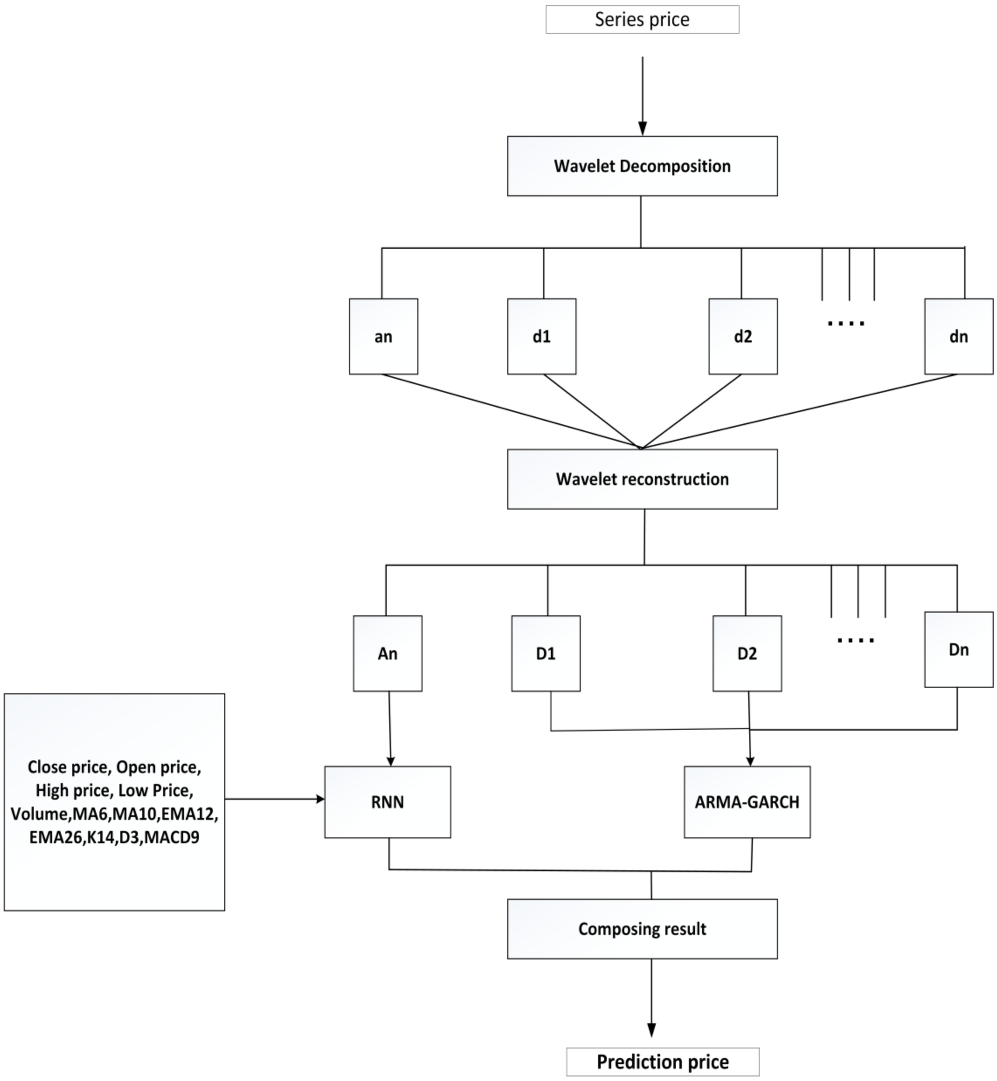

Fig. 3. proposed model structure

4. Results

To test the proposed model, the daily price series for 15 stocks from the New York Stock Exchange was collected from 24/6/2015 to 10/11/2017 (including 600 days) and its performance was compared with artificial neural network, ARIMA-GARCH, and ARIMA-ANN models for predicting the day-ahead and week-ahead. Figures 3 and 4 show the daily price charts of selected stocks in NYSE market. For the validation and comparison of the above models with the proposed model, Mean Absolute Percentage Error (MAPE) criterion was used. The method for calculating this criterion is given below.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_i - X_i|}{X_i} \tag{10}$$

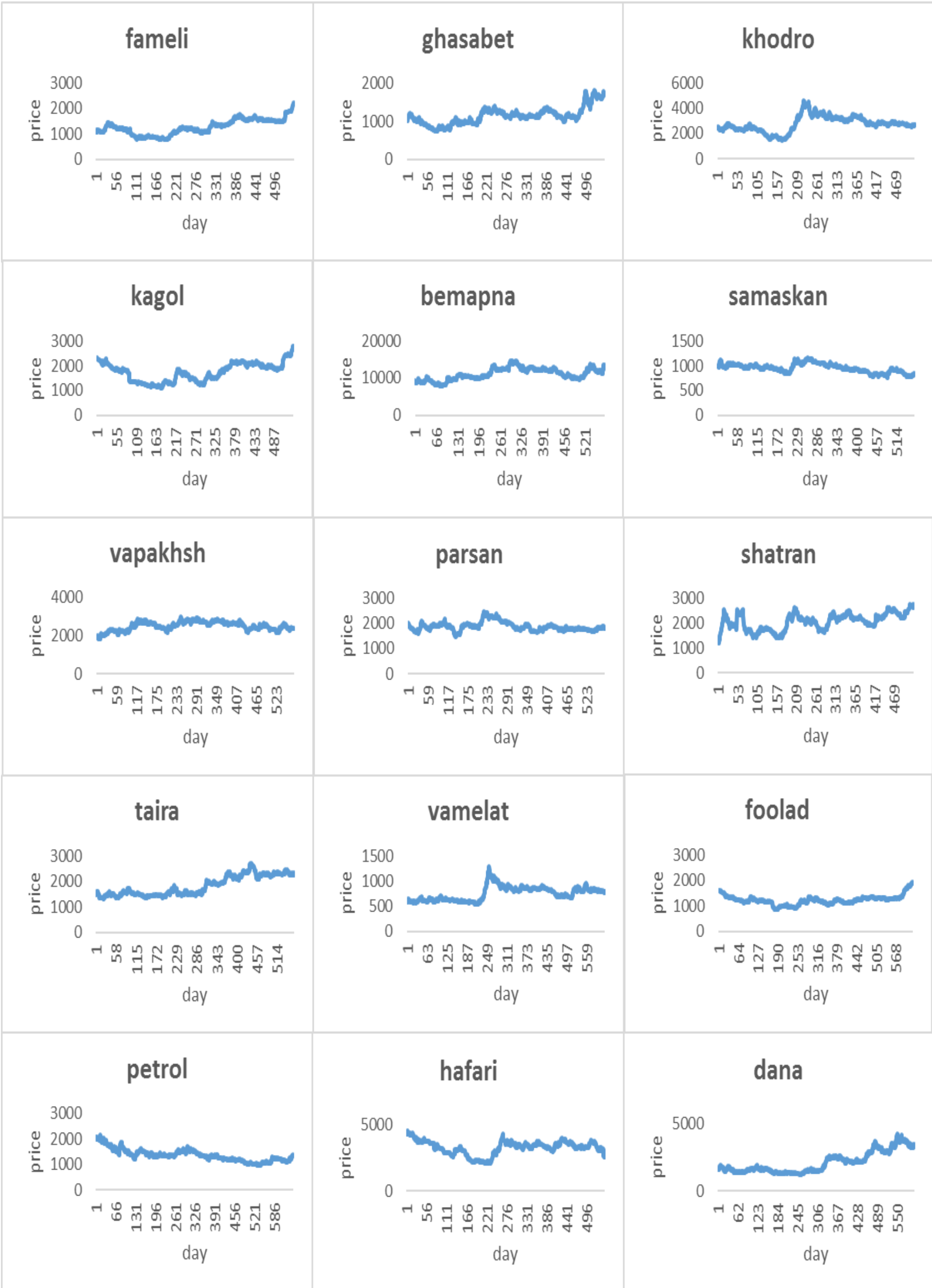

Fig. 4. daily price of New York Stock Exchange

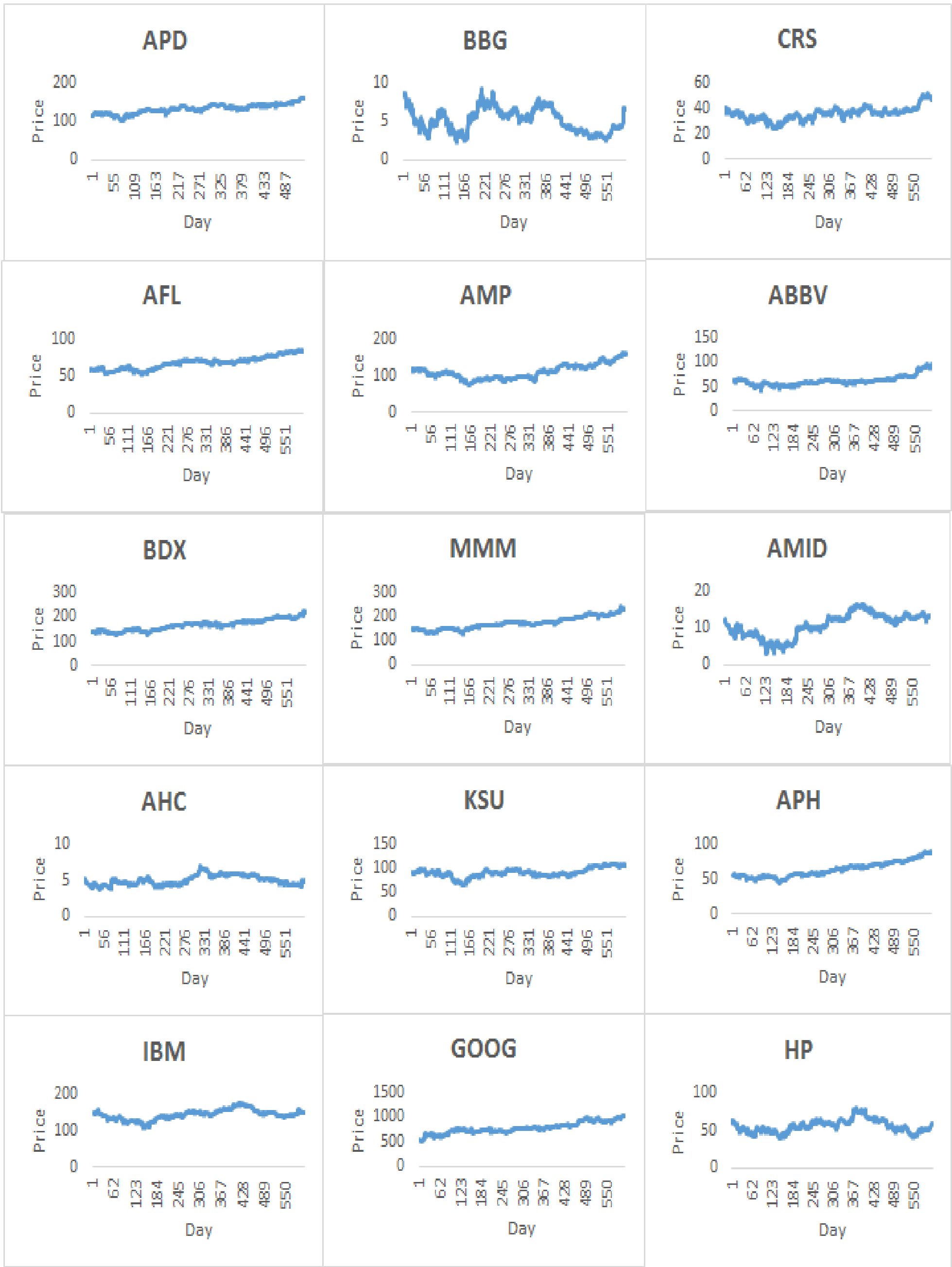

Fig. 5. daily price of New York Stock Exchange

Table 2. Error values of prediction based on MAPE criterion for 1-day prediction in New York Stock Exchange

	db2				db4				Haar				Other model		
	Level 2	Level 3	Level 4	Level 5	Level 2	Level 3	Level 4	Level 5	Level 2	Level 3	Level 4	Level 5	ANN	ARIMA-GARCH	ARIMA-ANN
APD	0.14	0.08	0.31	0.26	0.11	0.61	0.07	0.28	0.69	0.22	0.35	0.69	0.17	0.26	0.27
BBG	1.10	4.15	2.66	0.82	6.10	1.96	1.77	0.14	2.91	1.90	2.25	28.40	2.47	2.39	2.59
CRS	2.64	2.48	2.80	3.92	2.09	3.46	2.64	2.28	2.14	2.44	2.44	2.64	4.25	3.08	3.10
AFL	0.77	0.73	0.38	0.43	0.88	0.81	0.50	0.62	0.78	0.67	0.74	0.60	0.59	0.77	0.75
AMP	1.21	1.33	1.33	1.38	0.75	1.61	1.26	2.68	1.01	0.93	1.04	74.43	1.31	1.23	1.23
ABBV	0.84	0.52	0.45	0.49	0.30	0.13	0.99	2.86	0.05	0.15	0.00	0.83	0.23	0.36	0.33
BDX	1.82	2.12	1.12	0.80	0.96	1.09	1.26	1.11	0.99	1.39	1.00	0.83	1.99	2.40	2.37
MMM	0.16	0.09	0.45	0.29	0.04	0.28	0.41	0.52	0.03	0.15	0.18	1.65	1.38	0.44	0.51
AMID	1.49	1.35	0.25	0.11	1.56	1.77	0.05	1.29	0.49	1.32	1.24	1.44	0.42	0.65	0.78
AHC	1.92	0.33	2.73	2.38	3.18	0.59	0.21	32.7	2.23	2.20	3.46	5.26	2.12	1.96	2.05
KSU	0.31	0.47	0.75	0.35	0.02	0.83	0.66	6.73	0.16	0.19	0.21	0.20	0.35	0.53	0.52
APH	1.53	1.12	1.55	1.47	1.81	1.58	1.42	0.46	0.91	1.18	1.12	14816.2	1.58	1.89	1.85
IBM	0.10	0.68	0.49	0.01	1.24	0.20	0.80	1.44	0.07	0.69	0.99	22.02	0.73	0.78	0.77
Google	0.06	0.51	0.51	0.42	0.04	0.24	0.81	2.23	0.08	0.20	0.37	0.84	0.57	0.47	0.46
HP	3.70	2.28	0.58	1.64	1.73	0.12	0.34	2.49	2.02	1.86	0.14	8.46	0.96	1.16	1.16
Average	1.19	1.22	1.09	0.98	1.39	1.02	0.88	3.86	0.97	1.03	1.04	997.64	1.28	1.22	1.25

Table 3. Error values of prediction based on MAPE criterion for 1-day prediction in Tehran Stock Exchange

	db2				db4				Haar				Other model		
	Level 2	Level 3	Level 4	Level 5	Level 2	Level 3	Level 4	Level 5	Level 2	Level 3	Level 4	Level 5	ANN	ARIMA-GARCH	ARIMA-ANN
Fameli	0.23	0.50	0.08	0.21	27.94	1.75	0.24	1.43	1.81	1.52	1.16	0.07	2.12	1.87	3.20
Foolad	4.90	4.36	4.14	4.72	0.91	0.62	1.49	6.27	3.85	3.89	3.85	3.83	4.90	4.91	3.22
Ghasabet	1.49	1.43	1.61	2.31	0.57	0.37	0.07	1.94	1.91	1.69	1.35	3.09	1.44	1.23	1.44
Khodro	0.95	1.12	1.87	1.93	0.37	1.56	1.93	1.26	2.01	2.08	0.81	1.51	0.39	0.65	3.14
Kagol	1.98	2.03	1.41	0.50	2.88	1.31	1.77	0.49	1.09	0.90	0.59	1.90	0.81	0.88	2.10
Bemapna	2.25	2.49	2.50	2.63	6.06	2.65	3.00	2.64	2.24	1.66	1.94	2.20	2.97	3.17	3.34
Samaskan	0.41	0.12	0.05	0.32	5.64	1.32	0.92	0.09	0.63	0.09	0.04	0.22	0.44	0.40	0.18
Vapakhsh	0.01	0.24	0.62	0.19	1.21	0.23	0.49	1.35	0.46	0.20	0.01	0.69	0.38	0.36	0.24
Parsal	2.52	3.91	4.45	3.31	23.08	3.59	4.48	3.90	4.25	4.45	3.79	4.57	3.40	3.25	4.09
Petrol	0.31	0.71	0.25	0.90	3.84	0.08	0.79	0.55	7.41	0.99	1.03	1.44	0.74	0.54	0.43
Shatran	0.29	0.03	0.29	0.63	0.89	0.49	0.16	0.08	0.61	0.18	0.54	0.27	0.31	0.19	0.91
Taira	1.71	1.81	1.04	2.04	0.59	1.02	1.33	1.46	1.49	0.69	1.18	0.34	2.12	2.09	4.77
Vamelat	2.82	2.90	1.16	2.36	1.68	1.34	1.57	0.84	1.83	2.18	3.96	2.97	2.10	2.02	2.06
Hafari	1.25	0.90	0.17	1.59	1.65	1.11	0.97	0.64	5.62	5.16	6.07	4.66	1.54	1.56	2.18
Dana	6.38	4.37	5.32	5.43	2.03	4.10	5.05	3.12	4.74	4.79	4.48	3.81	4.66	4.37	2.84
Average	1.83	1.80	1.66	1.94	5.29	1.44	1.62	1.74	2.66	2.03	2.05	2.11	1.89	1.83	2.28

Table 4. Error values of prediction based on MAPE criterion for 5-day prediction in New York Stock Exchange

	db2				db4				Haar				Other model		
	Level 2	Level 3	Level 4	Level 5	Level 2	Level 3	Level 4	Level 5	Level 2	Level 3	Level 4	Level 5	ANN	ARIMA-GARCH	ARIMA-ANN
APD	0.38	0.79	0.79	0.35	0.96	0.61	1.35	0.70	0.46	0.35	0.41	0.29	5.48	0.56	0.51
BBG	3.55	6.66	3.46	7.47	4.19	11.78	2.60	0.93	3.29	6.76	7.54	6.17	2.31	3.48	1.46
CRS	7.08	7.57	1.97	1.08	4.27	5.79	1.90	6.34	5.94	3.15	1.80	1.27	3.14	2.13	2.22
AFL	0.60	0.82	0.46	0.43	0.52	0.93	0.50	0.45	0.48	0.44	0.57	0.46	0.40	0.33	0.35
AMP	1.99	1.02	1.17	3.24	1.21	0.48	0.65	0.66	1.32	1.58	1.34	1.83	2.71	0.63	0.65
ABBV	1.09	1.54	0.80	1.10	1.64	1.10	2.75	1.75	2.34	1.04	3.15	0.55	1.93	2.45	2.33
BDX	1.34	2.26	4.07	3.99	0.84	3.26	1.03	5.67	1.57	3.67	4.37	5.95	7.46	1.04	1.16
MMM	0.97	0.50	1.24	1.49	1.39	0.51	0.84	0.87	0.44	0.54	1.42	0.63	2.58	1.59	1.57
AMID	2.51	1.97	1.52	2.51	4.48	1.39	1.82	0.54	2.50	1.51	1.76	1.27	2.70	2.89	2.74
AHC	8.42	0.78	1.84	0.79	4.18	8.05	7.89	26.76	3.93	1.20	1.41	2.07	1.00	0.56	0.63
KSU	1.02	1.11	0.79	0.65	1.41	0.62	0.90	5.38	1.19	0.80	0.87	1.05	1.05	0.70	0.74
APH	0.56	0.78	0.90	1.92	0.65	0.60	0.58	0.40	0.69	0.69	0.41	1.35	0.94	0.48	0.47
IBM	0.30	0.40	1.95	2.89	0.43	1.35	0.42	1.57	0.89	2.32	1.95	0.48	0.28	0.33	0.28
Google	1.06	2.91	0.36	0.49	2.30	1.61	1.08	4.24	1.57	0.97	1.58	2.92	2.48	0.39	0.45
HP	5.63	7.51	6.71	6.16	7.28	6.08	7.12	3.45	6.79	6.94	7.88	38.52	6.73	6.45	6.90
Avergae	2.43	2.44	1.87	2.30	2.38	2.94	2.09	3.98	2.23	2.13	2.43	4.32	2.75	1.60	1.50

Table 5. Error values of prediction based on MAPE criterion for 5-day prediction in Tehran Stock Exchange

	db2				db4				Haar				Other model		
	Level 2	Level 3	Level 4	Level 5	Level 2	Level 3	Level 4	Level 5	Level 2	Level 3	Level 4	Level 5	ANN	ARIMA-GARCH	ARIMA-ANN
Fameli	11.03	7.29	7.97	10.63	7.12	7.06	9.79	8.98	13.50	9.92	10.37	10.94	6.54	6.71	9.60
Foolad	1.86	3.81	2.67	1.35	5.66	5.30	2.97	2.67	3.04	3.21	4.33	4.05	2.70	2.70	3.58
Ghasabet	2.88	1.55	1.69	3.22	1.74	1.35	1.14	2.39	3.79	2.89	2.39	2.59	1.60	1.51	1.76
Khodro	6.04	10.21	5.68	9.10	8.87	9.11	5.44	6.57	9.00	8.35	8.70	8.19	5.78	5.83	10.88
Kagol	6.33	6.39	6.56	5.56	5.98	5.94	4.09	4.42	5.74	7.16	7.75	7.11	4.19	4.26	4.22
Bemapna	1.12	1.84	1.71	2.28	4.40	3.55	2.22	2.43	1.66	2.13	2.92	2.36	2.51	2.51	2.03
Samaskan	0.78	0.58	0.65	0.57	15.02	0.66	0.53	0.32	5.50	0.42	0.29	0.40	0.28	0.32	0.33
Vapakhsh	3.30	3.78	1.80	3.21	4.97	2.63	1.68	4.43	3.03	3.00	5.09	3.77	2.75	2.74	2.15
Parsal	1.23	1.83	1.49	1.74	16.29	1.67	1.68	1.63	1.70	1.71	2.21	1.74	1.21	1.11	4.87
Petrol	1.48	1.99	1.52	2.12	5.06	1.12	1.13	1.48	1.21	1.03	1.20	2.36	1.96	1.99	1.69
Shatran	1.24	2.35	2.01	1.30	0.79	2.17	1.26	0.47	2.09	1.96	1.94	1.06	1.83	1.85	3.54
Taira	7.08	7.22	5.24	5.16	6.24	6.13	5.29	4.65	5.63	4.92	4.53	5.91	3.72	3.81	5.61
Vamelat	6.58	7.01	7.28	6.94	6.28	6.42	6.75	6.59	6.14	5.55	6.49	7.93	4.94	4.76	4.16
Hafari	17.31	19.21	19.49	20.60	16.66	13.96	18.72	19.18	19.59	15.98	17.46	18.29	17.86	17.68	19.45
Dana	5.30	2.02	2.04	2.04	2.42	2.23	2.10	2.08	2.24	3.21	2.86	2.37	3.38	2.71	2.24
Average	4.90	5.14	4.52	5.05	7.17	4.62	4.32	4.55	5.59	4.76	5.24	5.27	4.08	4.03	5.07

As shown in Table 2,3 and Figure 6,7, the performance of the proposed model by using different wavelet transform functions at most levels has shown better performance than the neural network, ARIMA-GARCH, and ARIMA-ANN models. Of course, in some cases, the levels of decomposition have a very high predictive error, which has caused the average prediction error of the other models to be very high. Among the models, the proposed model with Db4 wavelet transform at the decomposition level 4 has the best performance compared to other models.

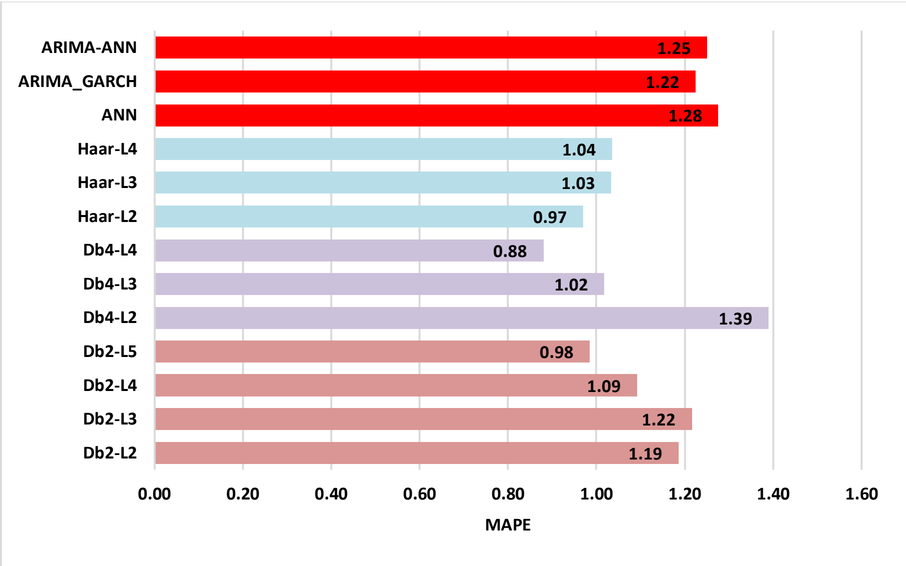

Fig. 6. Average of prediction errors of different models based on MAPE criterion for 1-day prediction in New York Stock Exchange

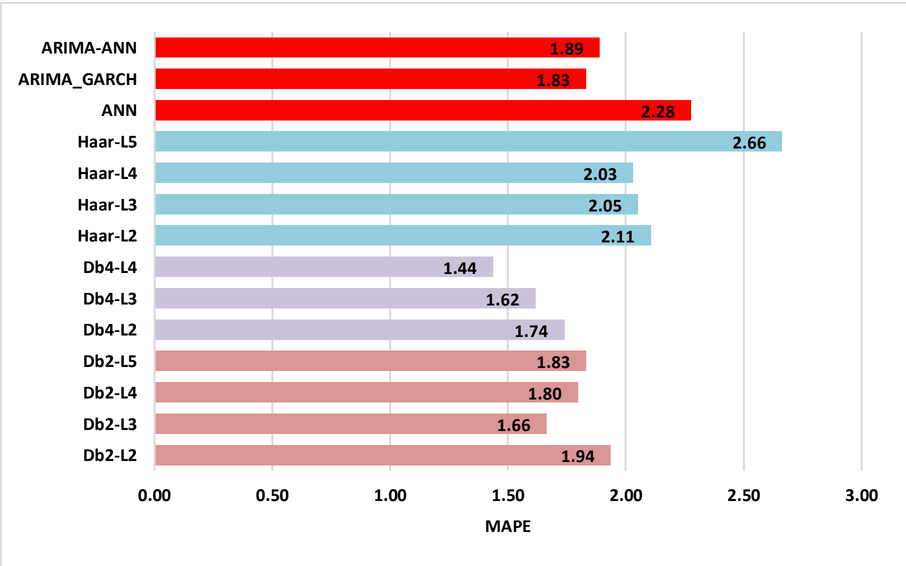

Fig. 7. Average of prediction errors of different models based on MAPE criterion for 1-day prediction in Tehran Stock Exchange

As shown in Table 4 and 5, the proposed model for 5-day prediction in most cases has a better performance than the neural network, but only the proposed model with the wavelet transform function combining the three functions Db2, Db4, and Haar at the decomposition level 3 and 4 have had the same performance of both ARIMA-GARCH and ARIMA-ANN models, and in the rest of the cases, the prediction error of the model is more than the other two models.

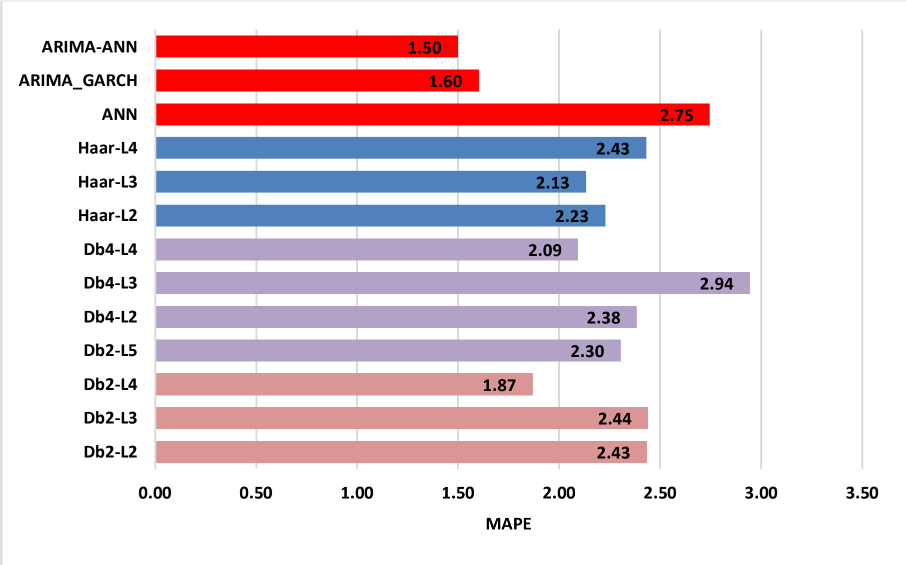

Fig. 8. Average of prediction errors of different models based on MAPE criterion for 5-day prediction in New York Stock Exchange

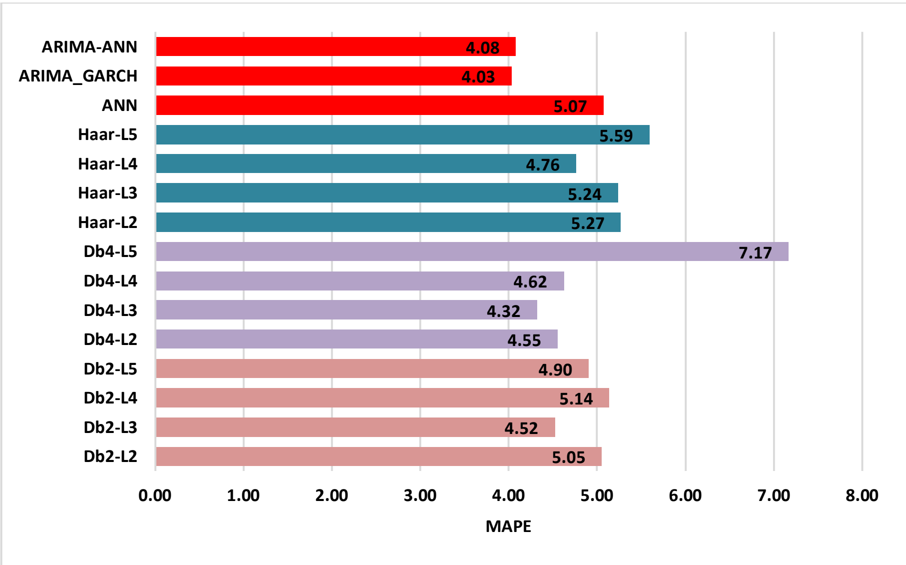

Fig. 9. Average of prediction errors of different models based on MAPE criterion for 5-day prediction in New York Stock Exchange

5. Discussion and Conclusions

Due to the nature of the financial time series fluctuations and the linear and nonlinear patterns in them, as well as the combination of models to consider these patterns and reduce their predictive errors, in this study, in the first step, using wavelet transform, the time series of prices is broken down, and then they will be reconstructed into an approximate (non-linear) series and some details (linear) series. ARMA-GARCH time series model is used to predict details time series and artificial neural network is used to predict approximate time series. In addition to the approximate series, a set of technical indicators will also be given as input to the neural network, and at the end, the results of the two models will be combined.

The results of proposed model implementation with different wavelet transform functions at two to five decomposition levels for the single-period forecast of the 15 stocks price of the Tehran and New York market showed that in most cases, this model performed better than three models of artificial neural network, ARIMA-GARCH, and ARIMA-ANN, and among them, the proposed model with the Db4 wavelet transform function at the decomposition level 4 has the best performance. also the result indicates the performance of forecasting models in New York stock exchange (mature market) is better than Tehran stock exchange (emerging market).

Also, the results showed that the proposed model would have better performance at the decomposition level 3 and 4. The results of the model implementation for the 5-day forecast showed that although the proposed model has a lower predictive error than the neural network, the ARIMA-GARCH and ARIMA-ANN models have less error than the other models.

References

1) Adhikari, R., & Agrawal, R. K. (2014). A combination of artificial neural network and random walk models for financial time series forecasting. Neural Computing and Applications, 24(6), 1441–1449.
2) Al Wadi, S., Ismail, M. T., Alkhahazaleh, M. H., & Karim, S. A. A. (2011). Selecting Wavelet Transforms Model in Forecasting Financial Time Series Data Based on ARIMA Model. Applied Mathematical Sciences, 5(7), 315–326.
3) Baba, N., & Kozaki, M. (1992). An intelligent forecasting system of stock price using neural networks. In [Proceedings 1992] IJCNN International Joint Conference on Neural Networks (Vol. 1, pp. 371–377 vol.1).
4) Babu, C. N., & Reddy, B. E. (2014). A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data. Applied Soft Computing, 23, 27–38.
5) Bowden, N., & Payne, J. E. (2008). Short term forecasting of electricity prices for MISO hubs: Evidence from ARIMA-EGARCH models. Energy Economics, 30(6), 3186–3197.
6) Chang, P. C., & Fan, C. Y. (2008). A hybrid system integrating a wavelet and TSK fuzzy rules for stock price forecasting. IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, 38(6), 802–815.
7) Chen, K. Y., & Wang, C. H. (2007). A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan. Expert Systems with Applications, 32(1), 254–264.
8) Chenoweth, T., & Obradović, Z. (1996). A multi-component nonlinear prediction system for the S&P 500 index. Neurocomputing, 10(3), 275–290.
9) Contreras, J., Espinola, R., Nogales, F. J., & Conejo, A. J. (2002). ARIMA Models to Predict Next-Day Electricity Prices. IEEE Power Engineering Review, 22(9), 57.
10) Dai, W., & Lu, C.-J. (2008). Financial Time Series Forecasting Using a Compound Model Based on Wavelet Frame and Support Vector Regression. 2008 Fourth International Conference on Natural Computation, 328–332.
11) Doulai, P., & Cahill, W. (2001). SHORT-TERM PRICE FORECASTING IN ELECTRIC ENERGY MARKET, 749–754.
12) Fernández-Rodríguez, F., González-Martel, C., & Sosvilla-Rivero, S. (2000). On the profitability of technical trading rules based on artificial neural networks:: Evidence from the Madrid stock market. Economics Letters, 69(1), 89–94.
13) Garcia, R. C., Contreras, J., Akkeren, M. van, & Garcia, J. B. C. (2005). A GARCH forecasting model to predict day-ahead electricity prices. IEEE Transactions on Power Systems, 20(2), 867–874.
14) Ghiassi, M., Saidane, H., & Zimbra, D. K. (2005). A dynamic artificial neural network model for forecasting time series events. International Journal of Forecasting, 21(2), 341–362.
15) Hyup Roh, T. (2007). Forecasting the volatility of stock price index. Expert Systems with Applications, 33(4), 916–922.
16) Jammazi, R., & Aloui, C. (2012). Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. Energy Economics, 34(3), 828–841.
17) Joo, T. W., & Kim, S. B. (2015). Time series

forecasting based on wavelet filtering. Expert Systems with Applications, 42(8), 3868–3874.
18) Kaastra, I., & Boyd, M. (1999). Designing a neural network for forecasting financial and economic time series. Neurocomputing, 10(3), 215–236.
19) Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time Series Forecasting Using Hybrid ARIMA and ANN Models Based on DWT Decomposition. Procedia Computer Science, 48(Iccc), 173–179.
20) Khashei, M., & Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. Applied Soft Computing Journal, 11(2), 2664–2675.
21) Lendasse, A., Bodt, E. de, Wertz, V., & Verleysen, M. (2000). Non-linear financial time series forecasting - Application to the Bel 20 stock market index. E.J.E.S.S., 14(1), 81–91. Retrieved from https://doi.org/10.1051/ejess:2000110
22) Pai, P.-F., & Lin, C.-S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. Omega, 33, 497–505.
23) Tan, Z., Zhang, J., Wang, J., & Xu, J. (2010). Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. Applied Energy, 87(11), 3606–3610.
24) Wang, J. Z., Wang, J. J., Zhang, Z. G., & Guo, S. P. (2011). Forecasting stock indices with back propagation neural network. Expert Systems with Applications, 38(11), 14346–14355.
25) Yu, L., Wang, S., & Lai, K. K. (2005). A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates. Computers and Operations Research, 32(10), 2523–2541.
26) Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159–175.