Wavelet Neural Network Prediction of Stock Performance

Jackson Criswell and En-Bing Lin*
Department of Mathematics
Central Michigan University
Mount Pleasant, Michigan 48859
USA
*Corresponding Author: lin1e@cmich.edu

Abstract

Wavelet methods and artificial neural networks are incorporated to examine the forecasting performance of the daily closing price of the Microsoft stock, NASDAQ:MSFT. An experimental analysis is performed to demonstrate improved performance of the wavelet neural network. Results in this study suggest that for these neurowavelet models a long history with a short training is ideal for stock prediction. This model could be used by investors, financial managers, or others to enhance their ability to select desired stocks.

1 Introduction

Accurate stock market prediction is not only an important topic in the understanding of the economy but also proves to be a challenging task for investors. It is hard to forecast trends in the stock market. This unpredictability in market trends in pervasive for many reasons. Not only is the future inherently unpredictable, but preferences of society can exhibit high spontaneity and sudden fluctuations in demand. In this work we perform wavelet analysis using the closing price of a stock as a time series signal. Consider price action as a time series and hypothesize that through the use of wavelets, data science, and machine learning we may accurately project how future price moves. Most time series exhibit a high degree of temporal and spatial dynamics and are described by some nonlinearity and singularities. Financial analysts attempt to provide rational explanations to problems that arise in the stock market. Through the development of theoretical models and predictive algorithms, effective information can be obtained for use in financial planning, analysis and management. Traditional time series models are limited with the assumption that data are stationary and thus unable to accurately capture nonlinearities or singularities in the signal. Stock market prices are usually the outcome of interaction among different nonlinear phenomena, which fluctuate across different spatial and temporal scales producing a chaotic response curve that is difficult to predict. Our purpose is to develop an adequate model that characterizes high complexity, and nonstationarity of the stock price as well as provide accurate forecasting.

Various approaches have been tested for forecasting and prediction of financial time series. The last couple of decades have seen remarkable progress in the ability to develop accurate time series models [1, 2, 5, 6]. Artificial neural networks (ANN) shows great strength in efficiency of curve-fitting and simulation through the diverse network designs and available training algorithms. ANN have become a popular tool among stock market and elsewhere in the realm of time series analysis [4, 8, 9]. However, when working with chaotic signals such as a series of daily stock price, a major bottleneck in the capabilities of ANN has been the ability to generalize a well-trained simulation to the accurate prediction of extreme events or extended forecasting. By including multi-resolution information from a wavelet decomposition of the time series as input the ANN model is specialized to form the hybrid wavelet neural network (WNN) or *neurowavelet* model. Among many models developed over this period, WNN type models have shown to be among the most promising in simulating stock market. This is in part due to the provable ability of ANN to efficiently approximate highly nonlinear relationships [3]. The power of ANN is combined with the efficient multiscale representation granted by the wavelet transform to increase forecast accuracy. This is exemplified in the demonstrable potential for robust prediction of nonlinear time series by models incorporating a neurowavelet technique [5, 8, 11, 12].

Wavelet analysis is a useful and powerful tool in performing time-frequency or time-space analysis of a time series. The wavelet transform can be used as a decomposition of a time series in to precise resolution in both time and frequency scale planes. Wavelets with multiresolution (MRA) properties have become useful tools in many applications, which include sub-band coding data compression, characteristic points detection, and noise reduction and others.

In this study, we combine wavelet analysis and artificial neural networks (ANN) as a hybrid stock market forecasting model WNN. By ANN we are referring to the concept of using a machine learning approach to perform stock market forecasting. In order to control for the innate uncertainty produced by neural network we introduce two controls in the implementation of the neurowavelet system. As a first measure, a genetic algorithm is used to only select the best trained networks, this process is repeated to generate a set of networks. Each network is used to create a prediction, and then these results are averaged together to give the reported forecast horizon. We use a neuro-wavelet method to perform forecasting on the Microsoft stock.

There are two main experimental variables and two hypothesis which were tested in this analysis. In the implemented neurowavelet system there are two ways to control the amount data being used for prediction, and these are the variables of interest here. The first is the length of the time series to be analyzed by the wavelet transform. The second experimental variable is the length of the training period supplied to WNN, which we refer to as the *Lookback time*. The tested research questions are as follows: Does the addition of additional history data to the wavelet MRA improve stock market prediction? How much look-back time is needed to cleanly reproduce market trends?

This paper is organized as follows. In Section 2, we briefly describe wavelet method, neural networks, data sources and comuter resources. We present our results in Section 3, followed by discussions in Section 4. We conclude with several comments and state our future plans in Section 5.

2 Methods

2.1 Wavelet Analysis

In what follows, we provide some background on wavelet analysis. This information includes a description of multiresolution analysis, scaling functions, wavelet functions, and the wavelet transform for both continuous and discrete signals.

A multiresolution analysis (MRA) [7, 10] consists of a sequence of successive approximation spaces $\{V_j\}_{j\in Z}$ of $L^2(R)$ with the following properties:

(i) $V_j \subset V_{j+1}$,

(ii) $\lim_{j\to\infty} V_j = \bigcup_{j\in Z} V_j$ is dense in $L^2(R)$,

(iii) $\bigcap_{j\in Z} V_j = \{0\}$,

(iv) $f(x) \in V_j \Longleftrightarrow f(2x) \in V_{j+1}$,

(v) $f(x) \in V_j \Longleftrightarrow f(x + 2^{-j}k) \in V_j, \ \ \forall\, k \in Z$,

(vi) There exists a function $\phi \in V_0$ so that $\{\phi(x-j)\}_{j\in Z}$ is an orthonormal basis of V_0.

ϕ is called a *scaling function* that generates a MRA with the above properties. Through translation and dilation of ϕ, a Riesz basis $\{\phi_{j,k}(x)\}_{k\in Z}$ is obtained for the subspace $V_j \subset L^2(R)$ by the properties (iv)(v), where

$$\phi_{j,k}(x) = 2^{\frac{j}{2}}\phi(2^j x - k), \quad j, k \in Z. \tag{1}$$

This family can be generally expressed as $\phi_{m,n}(x) = \frac{1}{a^{\frac{m}{2}}}\phi(\frac{x-nb}{a^m})$, for real numbers $a \neq 0$ and b.

Since $V_0 \subset V_1$, there is a set of coefficients $\{a_k\}_{k\in Z}$, so that ϕ satisfies the two–scale equation or refinement equation

$$\phi(x) = \sum_k a_k \phi(2x - k). \tag{2}$$

For every $j \in Z$, we define W_j to be the orthonormal complement of V_j in V_{j+1}, we then have

$$V_{j+1} = V_j \bigoplus W_j \tag{3}$$

and

$$W_j \perp W_{j'} \ \ if \ \ j \neq j'. \tag{4}$$

It follows that, for $j > J$

$$V_j = V_J \bigoplus (\bigoplus_{k=0}^{J-j+1} W_{J-k}). \tag{5}$$

By virtue of (ii) and (iii) above, this implies

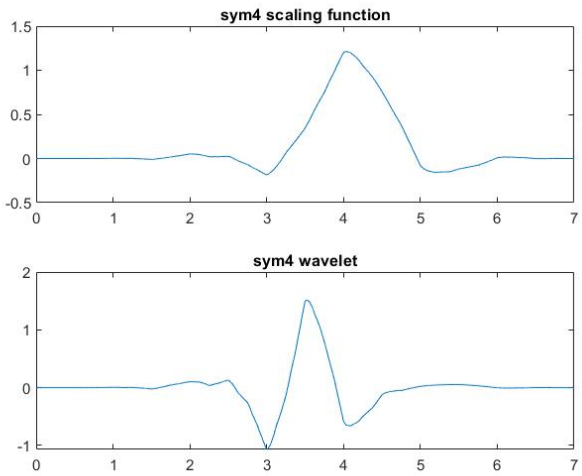

Figure 1: Symlet Scaling Function and Wavelet of Order 4

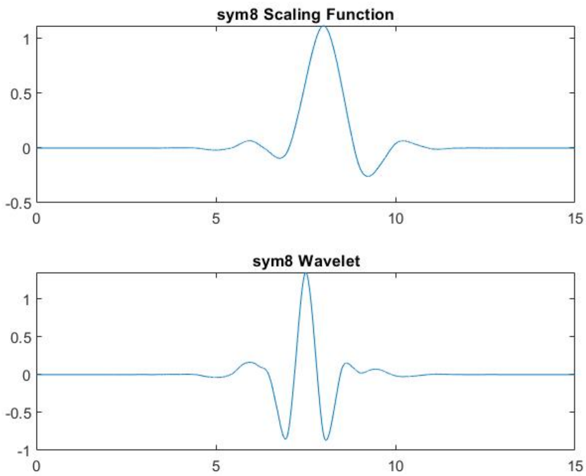

Figure 2: Symlet Scaling Function and Wavelet of Order 8

$$L^2(R) = \bigoplus_{j \in Z} W_j \tag{6}$$

which is a decomposition of $L^2(R)$ into mutually orthogonal subspaces. It turns out that a basis for W_0 can be obtained by dilating and translating a single function $\psi(x)$ called basic (mother) wavelet which is defined by (wavelet equation)

$$\psi(x) = \sum_k b_k \phi(2x - k) \tag{7}$$

where $b_k = (-1)^k a_{-k+1}$. In fact, $\{\psi_{j,k}(x) = 2^{\frac{j}{2}} \psi(2^j x - k)\}_{k \in Z}$ forms an orthonormal basis for W_j. Examples of scaling functions and wavelets are presented in Figures 1-4.

Let $P_j,\ Q_j$ denote the orthogonal projection $L^2 \to V_j,\ L^2 \to W_j$, respectively. Then

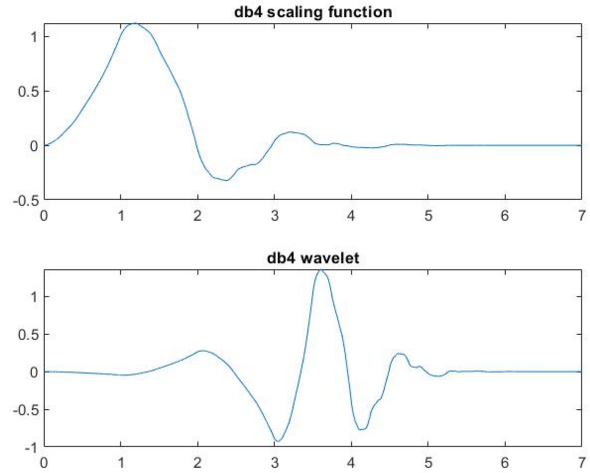

Figure 3: Daubechies Scaling Function and Wavelet of Order 4

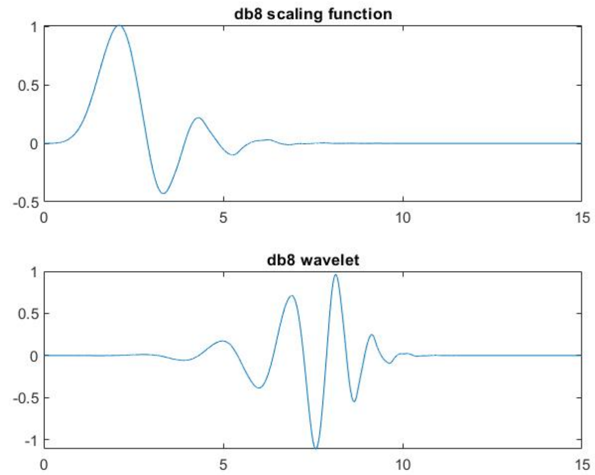

Figure 4: Scaling Function and Wavelet of Order 8

$$P_j f(x) = \sum_k \alpha_{j,k} \phi_{j,k}(x), \tag{8}$$

$$Q_j f(x) = \sum_k \beta_{j,k} \psi_{j,k}(x), \tag{9}$$

where the coefficients $\alpha_{j,k},\ \beta_{j,k}$ are given by the following inner products respectively:

$$\alpha_{j,k} =< f,\ \phi_{j,k} >= \int_{-\infty}^{\infty} f(x)\phi_{j,k}(x)dx, \tag{10}$$

$$\beta_{j,k} =< f,\ \psi_{j,k} >= \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx. \tag{11}$$

$P_j f$ converges to f in the L^2 norm which is the best approximation of f in V_j.

More precisely, the above coefficients can be obtained by applying wavelet transforms which are defined as follows.

The continuous wavelet transform is defined as:

$$[w_\psi x(t)](a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-b}{a}\right) dt \quad a > 0, b \in R, \tag{12}$$

where the symbol * represents the complex conjugate, $x(t)$ is the given signal and ψ is a wavelet. The discrete wavelet transform is defined as:

$$[Dw_\psi x(n)](a,b) = \sum_{n \in Z} x(n) g_{j,k}(n), \quad a = 2^j, b = k2^j, j \in N, k \in Z, \tag{13}$$

where $g_{j,k}$ are the coefficients of the wavelet equation associated with ψ.

To provide input data, we calculate wavelet coefficients by choosing appropriate wavelets and the level of MRA resolution.

2.2 Neural Networks

Artificial neural network (ANN) is one of the most promising methods in artificial intelligence. It is a mathematical method for data analysis based on learning and analyzing abilities. The idea is to simulate the human brain in the knowledge acquisition process to solve problems such as clustering, classification and prediction. ANNs set up components that possess essential properties of neurons and these are connected by specific weights. An ANN has input and output layers connected by a hidden layer. These successive layers receive the input information and propagate it towards the output layer. More precisely, the structure of the information processing system consists of a large number of interconnected processing elements (neurons). The networks learn the systems by adjusting to the synaptic connections that exist between neurons. The nonlinear autoregressive neural network is one of the basic models of ANN appropriate for estimation of future values of the input variable. It performs multistep neural network prediction which is for multi-step ahead prediction. In this context, dynamic networks with feedback can be transformed between open-loop and closed-loop modes. Closed-loop networks continue to predict when external feedback is missing, by using internal feedback.

There are many training functions used to train an ANN. The idea is to forecast future values of a time series, based on its historical (previous) values, utilizing additional external time series with some time delay parameters. The network training is performed, by some back propagation algorithm, and uses steepest descent method [8] to obtain least error between the real data and the predicted values.

2.3 Wavelet Neural Network

The idea of wavelet neural network (WNN) is to combine wavelet analysis and ANN as a hybrid time series forecasting model. By ANN we are referring to the concept of using a machine learning approach to perform forecasting. We first use wavelet tools to obtain wavelet coefficients of input data, we then use them to process through ANN which is one of the most powerful and useful methods in artificial intelligence. ANN is a typical mathematical method for data analysis based on learning and analyzing abilities so that we simulate the data through the knowledge acquisition process to solve problems as we described in the previous sub-session. The incorporation of wavelet analysis and ANN gives rise to more accurate, efficient and effective ways of classification and forecasting. The robustness and flexibility of our WNN system have been proven through additional experiments on other data. In addition, our model has been extended to incorporate multifarious extraneous measured input signals, without loss of algorithmic efficiency.

2.4 Data Sources

Microsoft (MSFT) price can be found at https://www.macrotrends.net/ which includes 35 years stock prices and other financial information.

2.5 Computer Resources

All computational experiments were performed using the High Performance Computing Center (HPCC) at Michigan State University's Institute for Cyber Enabled Research (ICER). The calculations were run on a single machine equipped with a AMD EPYC 7H12 64-Core Processor @ 2.6GHz and 996 GB of DDR4 ECC RAM. By way of this configuration a homogeneous parallel cluster was generated using 101 logical cores with 4 GB of memory allocated to each core and used for each experimental trial.

The predictive neurowavelet system was implemented as a MATLAB function. Computations were performed using MATLAB/2021a and the following toolboxes: Deep Learning, Parallel Computing, and Wavelet.

3 Results

The data was divided as 52% for training, 15% for validation and the final 33% as testing data.

After extensive testing with 100,500,2500, and 7500 days of lookback superior results were observed with 100 days of training. Based on this it was determined that a more accurate stock prediction is determined using less days WNN training. This is an unexpected result and stands in contrast of previous experiments applying the same system to a more natural signal

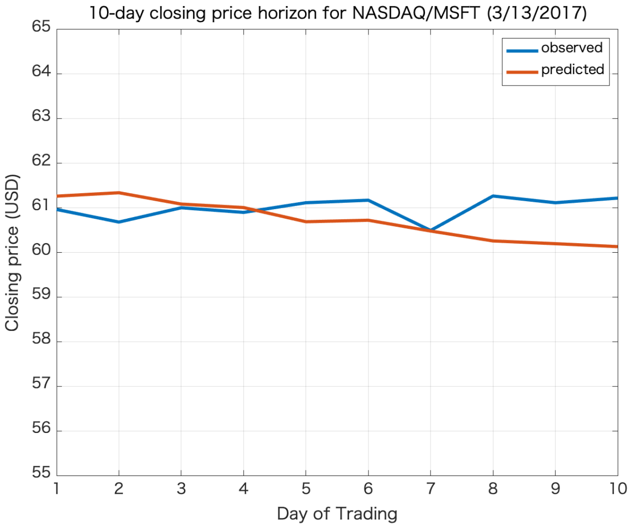

Figure 5: 10 day forecast horizon (based on the data since 1986)

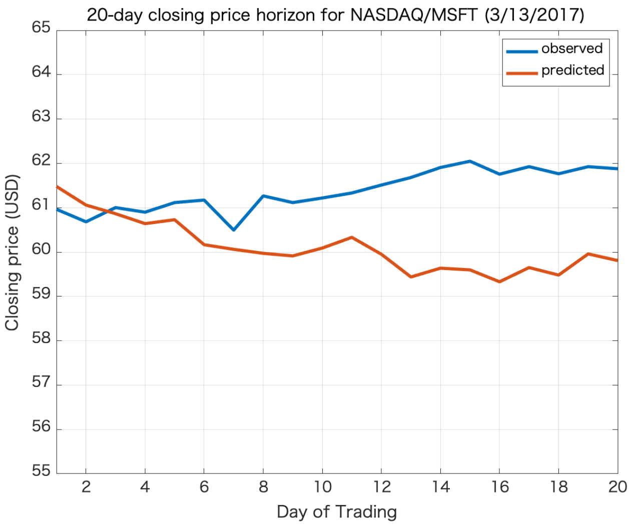

Figure 6: 20 day forecast horizon (based on the data since 1986)

[1]. Based on further detailed tuning, it was determined that 66 days of lookback should be used as the optimal length of history for the target signal.

To perform the WNN process, we use 6 delays, one hidden layer of 13 neurons, level 13 DWT with sym3 wavelet. In order to get reasonably good results, we run 100 trials and took the average. Based on the data since 1986, we obtained 10-day, 20-day, and 30-day closing price forecasting for NASDAQ/MSFT as shown in Figures **??**, **??**, and 7. In a similar way, we based on the data since 1995, we obtained and presented the results in Figures 8, 9, and 10. The model performance is measured by Mean Squared Error (MSE), Root Mean Square Error(RMSE), RMS Relative Error (RMSRE), and Mean Absolute Percent Error as defined in Table 3. We calculated the performance of the above methods and present the results in Table 3 and Table 4. The steps of the algorithm are presented in the appendix.

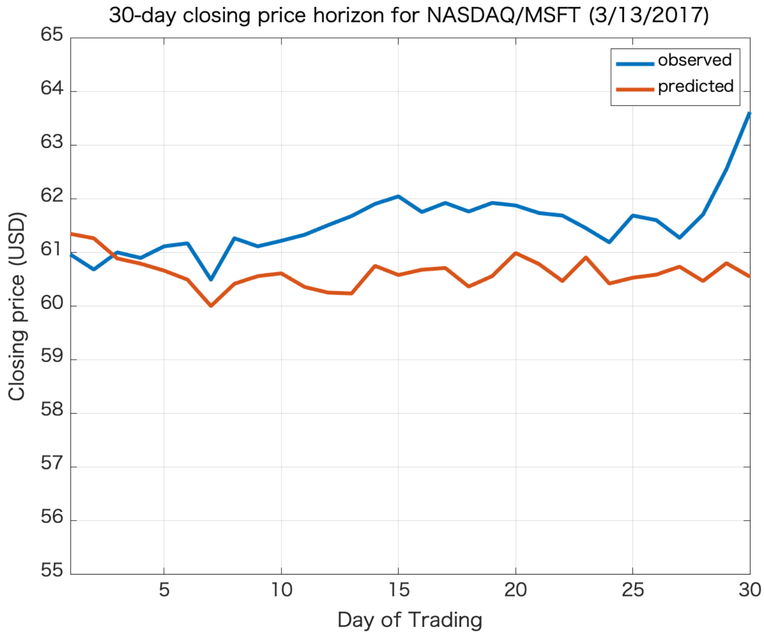

Figure 7: 30 day forecast horizon (based on the data since 1986)

Table 1: Performance Statistics

MSE	$\frac{1}{n}\sum_{t=1}^{n} e_t^2$
RMSE	$\sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$
RMSRE	$\sqrt{\frac{1}{n}\sum_{t=1}^{n} (\frac{e_t}{y_t})^2}$
MAPE	$\frac{1}{n}\sum_{t=1}^{n} \lvert\frac{e_t}{y_t}\rvert$

Table 2: Model Configurations (Hyperparameters)

Model	Hidden Layers	Wavelet Family	Resolution	Buffer	Lookback	History
WNN86	(13)	*Sym3*	13	6	66	7814
WNN95	(13)	*Sym3*	13	6	66	5539
MLP[4]	(70,28,14,7)	-	-	1	-	2750
UA [4]	(70,70)	-	-	-	-	2750

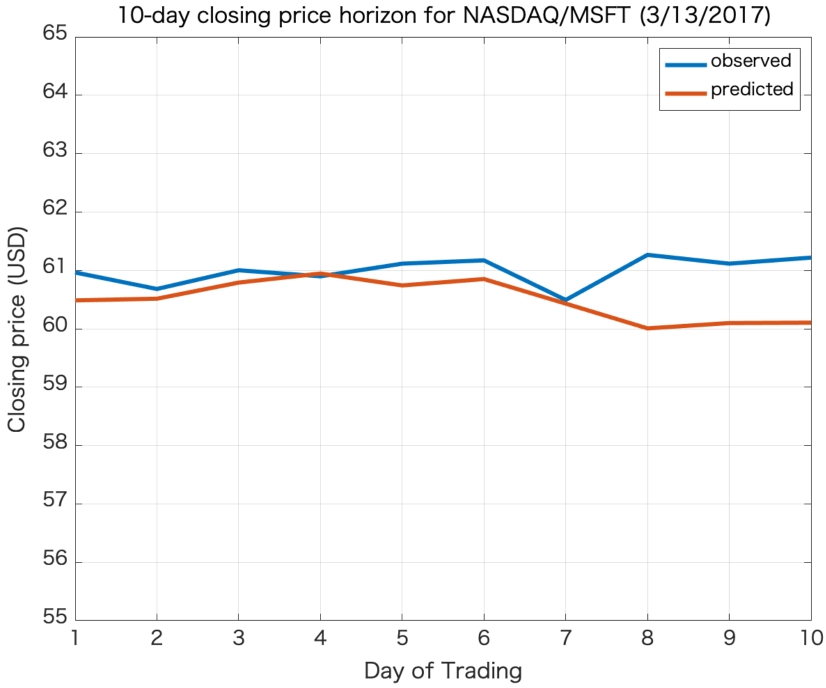

Figure 8: 10 day forecast horizon (based on the data since 1995)

4 Discussions

The performance with longer history data shows better than less history data. Remarkably, this means forecasting performs better if we are able to supply a longer history to the WNN during the closed-loop training and testing segments of model prediction. However, in calibrating the length of lookback it was determined that a shorter training data produces the best result. Together, this seems to imply that a long history with a short training is ideal for stock prediction.

Our WNN method shows better results than some neural network method [4, 12]. Figure 11 presents predictive results of a final response showing an improvement by the WNN presented here over the neurowavelet designs developed in [12]. The details of the hyperparameters representing some comparable models from [4] are shown in Table 3. Table 5 provides an explicit comparison of the measured statistics further demonstrating the superiority of the WNN method.

The performance gain seen with WNN method is likely due to the higher level wavelet decomposition incorporating more information of the time series. Also, WNN method can do better jobs on non-stationary time series and allows choosing different levels of resolution.

5 Conclusions

We use wavelet method together with ANN in predicting stock price and see the effects of our method by comparing various errors. Wavelet analysis has the ability to improve forecasting by capturing useful information on various resolution levels of a signal. ANN is very useful in

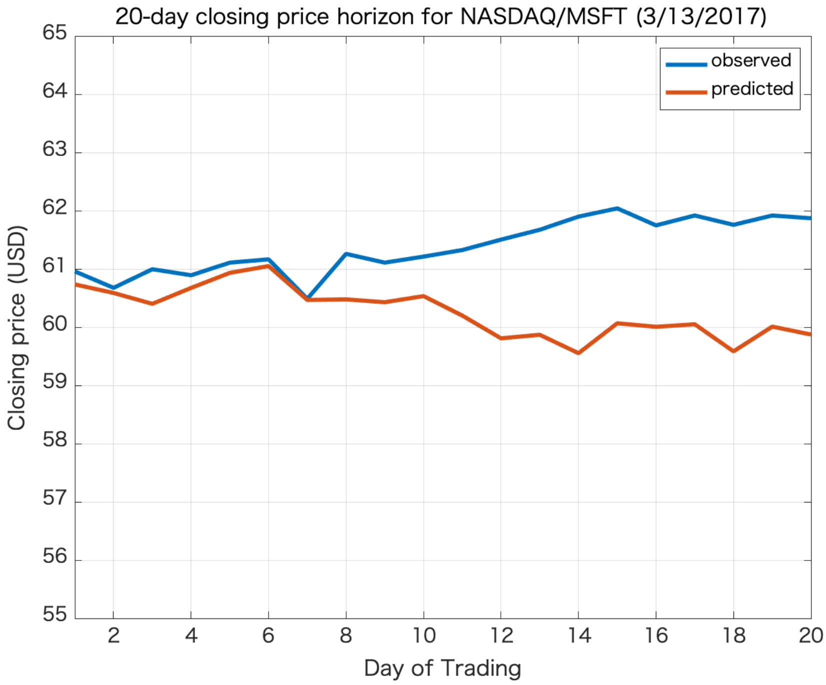

Figure 9: 20 day forecast horizon (based on the data since 1995)

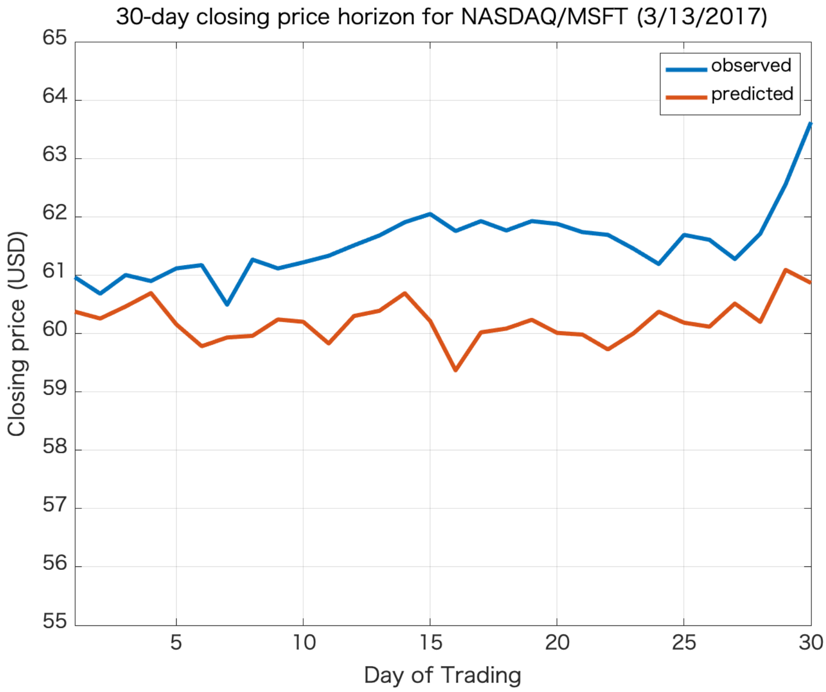

Figure 10: 30 day forecast horizon (based on the data since 1995)

modeling and forecasting time series. Combining both techniques, we obtain better results than other methods. Overall, WNN model using longer training produces better results than other methods. In fact, it turns out that short-term traders are usually better served by waiting for

Table 3: Forecast Performance (history: 3/13/1986-3/10/2017)

starting 3/13/2017		
horizon (days)	***MSE***	***RMSRE***
10	0.40	0.00045
20	2.52	0.0010
30	1.27	0.00050

Table 4: Forecast Performance (history: 3/13/1995-3/10/2017)

starting 3/13/2017		
horizon (days)	***MSE***	***RMSRE***
10	0.44	0.00081
20	1.89	0.00091
30	2.10	0.00072

Table 5: Comparison of Method Performance

Model	RMSE	MAPE
WNN86-10	0.6296	0.00045
WNN95-10	0.6642	0.00082
MLP-SP500[4]	44.5137	0.0118
UA-SP500[4]	25.4851	0.0067

confirmation of an output at hand, rather than trying to predict what an output will be in the long run. From what is learned of the outcome, traders establish significant stages to buy or sell that should be based on what price is actually doing, rather than what we expect it to do.

We will explore various tasks related to the method we presented in this paper, such as the level of wavelet resolution, genetic evolution parameters, number of neurons, size of delay buffer, and the structure of the hidden layer. We will also perform forecasting for other stocks. We envision our study will have impact on understanding and predicting the trends crucial in maintaining a stable marketplace environment with capacity for safe, effective, and profitable

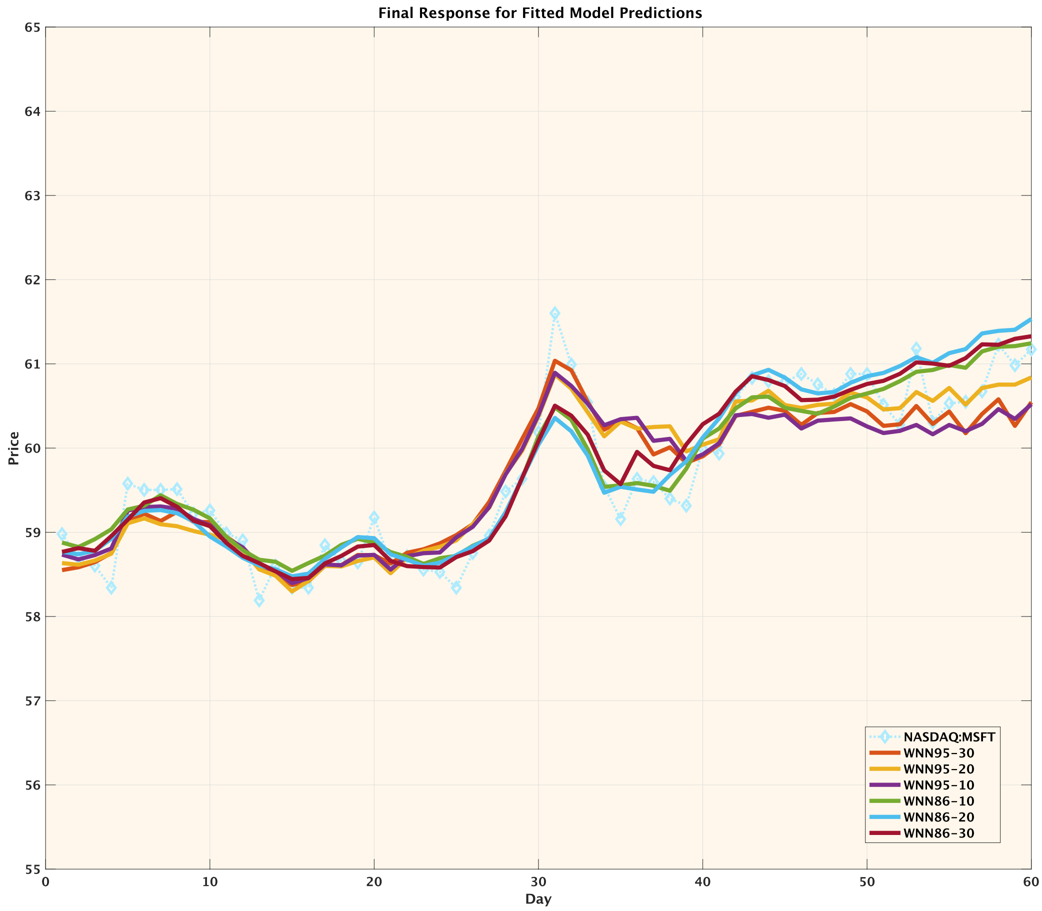

Figure 11: Comparison of prediction to fitted testing data for each WNN method.

stock trading. In this way we hope to produce work that helps improve the international economy and the well-being of our society.

Appendix

The complexity and power of ANN is achieved by the interaction of several neurons through the nonlinear process. The algorithm is based on the following Scaled Conjugate Gradient (SCG) training algorithm which is briefly described as follows. The idea is basically to minimize the error function $E(w)$ of the weight vector w with the following steps.

Step 1. Select the initial weight vector w_1 and let k=1.

Step 2. Determine a search directionp_k and a step size h_k so that $E(w_k + h_k p_k) < E(w_k)$.

Step 3. Update vector $w_{k+1} = w_k + h_k p_k$.

Step 4. If $E'(w_k) \neq 0$ then set $k = k + 1$ and go to Step 2 else return w_{k+1} as the desired minimum.

Acknowledgment

The authors wish to thank anonymous referees' comments which improve the manuscript.

References

[1] J. A. Criswell, and E. B. Lin, *River Flow Forecasting Using An Inverse Wavelet Transform Neural Network Approach,* Int. J. of Applied Mathematics, Computational Science and Systems Engineering 3, 67-70, 2021.

[2] J. A. Criswell, and E. B. Lin, *Noise Removal in River Flow Forecasting,* Proceedings of 2021 23rd International Conference on Process Control (IEEE), 219-224, June 2021.

[3] G. Cybenko, Approximation by superpositions of a sigmoidal function, Mathematics of control, signals and systems, 2(4), 303-314. 1989.

[4] P. Gao, R. Zhang, and X. Yang, *The Application of Stock Index Price Prediction with Neural Network,* Mathematical and Computational Application, 25, 53, 2020.

[5] Hsieh, T. J., H. F. Hsiao, and W. C. Yeh, *Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm,* Applied soft computing 11, 2510-2525, 2011.

[6] E. B. Lin, and J. Criswell, *On Wavelet Neural Networks and River Flow Forecasting,* Proceedings of 2021 6th International Conference on Mathematics and Artificial Intelligence, 20-24, 2021.

[7] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[8] Ortega, Luis F, *A neuro-wavelet method for the forecasting of financial time series*, Proceedings of the World Congress on Engineering and Computer Science, 24-26, 2012.

[9] Popoola, Ademola Olayemi, *Fuzzy-wavelet method for time series analysis*, University of Surrey, United Kingdom, 2006.

[10] Van Fleet, P. J. (2008), Discrete Wavelet Transformations An Elementary Approach with Applications, Hoboken, NJ, John Wiley and Sons.

[11] Kai-Cheng Wang, Chi-I Yang1 and Kuei-Fang Chang, *Stock prices forecasting based on wavelet neural networks with PSO,* MATEC Web of Conferences 119, 01029, 2017.

[12] Tianxing Zheng, Kamaladdin Fataliyev, Lipo Wang, *Wavelet neural networks for stock trading,* Proceedings of SPIE - The International Society for Optical Engineering, 2013.