



Wavelet low- and high-frequency components as features for predicting stock prices with backpropagation neural networks

Salim Lahmiri *

ESCA School of Management, 7, Abou Youssef El Kindy Street, BD Moulay Youssef, Casablanca, Morocco

Received 21 January 2013; revised 13 September 2013; accepted 4 December 2013
Available online 14 December 2013

KEYWORDS

Stock prices;
Discrete wavelet transform;
Approximation and detail coefficients;
Backpropagation neural networks;
Forecasting

Abstract This paper presents a forecasting model that integrates the discrete wavelet transform (DWT) and backpropagation neural networks (BPNN) for predicting financial time series. The presented model first uses the DWT to decompose the financial time series data. Then, the obtained approximation (low-frequency) and detail (high-frequency) components after decomposition of the original time series are used as input variables to forecast future stock prices. Indeed, while high-frequency components can capture discontinuities, ruptures and singularities in the original data, low-frequency components characterize the coarse structure of the data, to identify the long-term trends in the original data. As a result, high-frequency components act as a complementary part of low-frequency components. The model was applied to seven datasets. For all of the datasets, accuracy measures showed that the presented model outperforms a conventional model that uses only low-frequency components. In addition, the presented model outperforms both the well-known auto-regressive moving-average (ARMA) model and the random walk (RW) process.

© 2013 Production and hosting by Elsevier B.V. on behalf of King Saud University.

1. Introduction

Forecasting stock markets has long been investigated by researchers and professionals. Indeed, a large number of computing methods for stock prediction have been proposed in the literature (see [Atsalakis and Valavanis \(2009\)](#) and [Bahrammirzaee \(2010\)](#) for surveys). However, due to non-stationary, high volatility clustering and chaotic properties of the stock market prices, the prediction of share prices is always considered to be a difficult and challenging task. Recently, multi-resolution techniques such as the wavelet transform ([Mallat, 1989](#); [Daubechies, 1992](#)) have been successfully applied to both engineering problems ([De and Sil, 2012](#); [Rikli,](#)

* Address: Department of Computer Science, University of Quebec at Montreal, 201 President-Kennedy, Local PK-4150, Montreal, Québec H2X 3Y7, Canada. Tel.: +1 5149876516.
E-mail addresses: slahmiri@esca.ma, lahmiri.salim@courrier.uqam.ca.

Peer review under responsibility of King Saud University.



Production and hosting by Elsevier

2012) and financial time series studies (Li et al., 2006; Huang and Wu, 2008, 2010; Hsieh et al., 2011; Huang, 2011; Wang et al., 2011; Kao et al., 2013; Lahmiri, 2013) because of its powerful feature extraction capability. The wavelet transform is a signal processing technique that simultaneously analyzes the time domain and the frequency domain. In particular, the wavelet transform decomposes a time series into subsequences at different resolution scales. In particular, it decomposes given data into high and low-frequency components. At high frequency (shorter time intervals), the wavelets can capture discontinuities, ruptures and singularities in the original data. At low frequency (longer time intervals), the wavelet characterizes the coarse structure of the data to identify the long-term trends. Thus, the wavelet analysis allows us to extract the hidden and significant temporal features of the original data.

Li et al. (2006) applied the discrete wavelet transform to decompose the Dow Jones Industrial Average (DJIA) index time series and to extract features derived from approximation coefficients such as energy, entropy, curve length, non-linear energy and other statistical features. Finally, a genetic programming algorithm was used for forecasting purposes. They concluded that the wavelet analysis provides promising indicators and helps to improve the forecasting performance of the genetic programming algorithm. Huang and Wu (2010) used a discrete wavelet transform to analyze financial time series, including the National Association of Securities Dealers Automated Quotations (NASDAQ, United States), Standard & Poors 500 (S & P 500, United States), Cotation Assistée en Continu (CAC40, France), Financial Times Stock Exchange (FTSE100, United Kingdom), Deutscher Aktienindex (DAX30, Germany), Milano Italia Borsa (MIB40, Italy), Toronto Stock Exchange (TSX60, Canada), Nikkei (NK225, Japan), Taiwan Stock Exchange Weighted Index (TWSI, Taiwan) and the Korea Composite Stock Price Index (KOSPI, South Korea). A Recurrent Self-Organizing Map (RSOM) neural network was used for partitioning and storing the temporal context of the feature space. Finally, a multiple kernel partial least squares regression was used for forecasting purposes. The simulation results indicated that the presented model achieved the lowest root-mean-squared forecasting errors in comparison with neural networks, support vector machines or the traditional general autoregressive conditional heteroskedasticity (GARCH) model. Hsieh et al. (2011) applied the wavelet decomposition to analyze the stock price time series of the Dow Jones Industrial Average Index (DJIA), London FTSE-100 Index (FTSE), Tokyo Nikkei-225 Index (Nikkei) and Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX). Then, they used a recurrent neural network (RNN) to perform the forecasting task. The Artificial Bee Colony algorithm (ABC) was adopted to optimize the RNN weights and biases. The authors concluded that the proposed system is highly promising based on the obtained simulation results. Huang (2011) combined wavelet analysis with kernel partial least square (PLS) regression for stock index forecasting, including the NASDAQ (US), S & P 500 (US), TSX60 (Canada), NK225 (Japan), TWSI (Taiwan), and KOSPI (South Korea), CAC40 (France), FTSE100 (UK), DAX30 (Germany) and the MIB40 (Italy). The DWT was employed to identify financial time series characteristics, and the PLS was used to create the most efficient subspace that maintains maximum covariance between inputs and outputs. In terms

of the forecasting errors, the empirical results showed that the DWT-PLS model outperformed traditional neural networks, support vector machines and GARCH models. Wang et al. (2011) used wavelets to transform the Shanghai Stock Exchange (SCE) prices into multiple levels of decomposition. Then, for each level of decomposition, the backpropagation neural network (BPNN) was adopted to predict SCE prices while using low-frequency coefficients. The authors found that the BPNN with fourth decomposition level low-frequency coefficients outperforms a BPNN that uses past values of the original data. Lahmiri (2013) applied discrete wavelets to decompose the S & P 500 price index. The low-frequency coefficient time series were extracted, and out-of-sample predictions of the S & P 500 trends were conducted. Support vector machines (SVM) with different kernels and parameters were used as the baseline forecasting model. The simulation results reveal that the SVM with the wavelet analysis approach outperforms the SVM with macroeconomic variables or technical indicators as predictive variables. The author concluded that the wavelet transform is appropriate for capturing the S & P 500 trend dynamics.

To predict future stock prices, previous studies have used only approximation coefficients in an attempt to work with de-noised data. However, working with approximation decomposition coefficients is useful only in capturing major trends in the data. Indeed, approximation coefficients capture major trends of a time series whereas detail coefficients capture only deviations in the time series. As a result, choosing approximation coefficients as predictive inputs is not appropriate when capturing the overall characteristics of the original data. To take full advantage of the wavelet transforms, detail coefficients should also be used as predictors of future stock prices because detail coefficients are suitable for detecting local hidden information, such as abrupt changes, outliers and short discontinuities in stock prices. We argue that these features could improve the forecasting accuracy of machine learning methods.

In summary, wavelet transforms decompose a signal at different dilations, to obtain those approximation coefficients that represent the high-scale and low-frequency components and the detail coefficients that represent the low-scale and high-frequency components. From the viewpoint of feature extraction, high-frequency components are a complementary part of low-frequency components; in this way, they can capture the missing features that frequency components do not capture. Combining the two frequency components (two types of features) could provide better accuracy in the prediction of future stock prices.

To examine the effectiveness of high-frequency coefficients obtained from wavelet transforms in the prediction of stock prices, artificial neural networks (NN) are adopted in this study as the main machine learning approach for forecasting prices. Indeed, they are very popular as nonlinear stock market forecasting models because the behavior of share prices is nonlinear (Atsalakis and Valavanis, 2009; Bahrammirzaee, 2010; Wang et al., 2011). Artificial neural networks are nonlinear methods that can learn from patterns and capture hidden functional relationships in given data even if the functional relationships are not known or are difficult to identify (Zhang et al., 1998). In particular, they are capable of parallel processing information when there is no prior assumption about the model form. In addition, artificial neural networks are adap-

tive; hence, a network is capable of modeling non-stationary and dynamic data and systems. Finally, when motivated by short-term information arrivals, high-frequency financial data that is observed every successive minute are used in this work for two major reasons (Engle and Russell, 2010): intraday (high-frequency) financial data typically contain strong periodic patterns, and unlike their lower frequency counterparts (for example, daily and weekly data), high-frequency financial data typically display strong dependence. Therefore, it could be encouraging to predict such data. Indeed, because high-frequency (minute-by-minute) financial data are characterized by periodicities and strong dependences (Engle and Russell, 2010), these two important patterns could facilitate predicting the behavior of the stock market. We would like to note that this type of analysis, specifically using wavelet analysis high-frequency components to obtain an improvement in the accuracy of predicting stock prices while using minute-by-minute observed data, has not been conducted before, to the best of our knowledge.

In summary, the contribution of our work is as follows. First, both low- and high-frequency components are used as input variables to forecast future stock prices using backpropagation neural networks (BPNN). Second, the proposed model will be compared to the standard approach, in which only low-frequency components are used as predictive inputs to be fed into the BPNN. Third, to assess the effectiveness of our model against statistical models, its performance will be compared to that of the well-known auto-regressive moving-average (ARMA) model, which is a popular statistical approach for financial time series forecasting (Denton, 1995; Hann and Steurer, 1996; Taskaya and Casey, 2005; Rout et al., 2014). In addition, the prediction accuracy of our model is also compared to that of a random walk (RW) process. For example, if stock prices follow a random walk, then they cannot be predictable according to the efficient market hypothesis (Fama, 1965). Thus, comparing our model with the random walk process allows us to check whether our approach can be effective in stock market prediction from a financial theoretical point of view. Fourth, forecasting minute-by-minute financial data by virtue of DWTs and BPNNs is considered in our study for the first time.

The remainder of this paper is organized as follows. In Section 2, the design of the proposed prediction system is provided. Section 3 presents the ARMA model, the random walk process and performance measures. The empirical results

are presented in Section 4. Finally, Section 5 concludes the paper.

2. Design of the proposed prediction system

The proposed automated stock price prediction system consists of three steps: (1) The original stock price time series $s(t)$ are processed with a discrete wavelet transform (DWT); (2) both approximation $a(t)$ and detail $d(t)$ coefficients are extracted to form the main feature vector that characterizes the original time series; and (3) the resulting feature vector feeds the input of a back-propagation neural network (BPNN). The design of the proposed system is shown in Fig. 1. The wavelet transform, artificial neural networks, and performance measures are described in more detail next. For comparison purposes, a similar prediction system is simulated in which detail coefficients are excluded as predictors.

2.1. The wavelet analysis

In this section, a brief description of the wavelet transform is given. A thorough review of the wavelet transform is provided in Mallat (1989) and Daubechies (1990, 1992). The wavelet analysis is a mathematical method that allows decomposing a given signal $s(t)$ into many frequency bands or at many scales. In particular, the signal $s(t)$ is decomposed into smooth coefficients a and detail coefficients d , which are given by

$$a_{j,k} = \int s(t)\Phi_{j,k}(t)dt \quad (1)$$

$$d_{j,k} = \int s(t)\Psi_{j,k}(t)dt \quad (2)$$

where Φ and Ψ are, respectively, the father and mother wavelets, and j and k are, respectively, the scaling and translation parameters. The father wavelet approximates the smooth (low-frequency) components of the signal, and the mother wavelet approximates the detail (high-frequency) components. The father wavelet Φ and the mother wavelet Ψ are defined as follows:

$$\Phi_{j,k}(t) = 2^{-j/2}\Phi(2^{-j}t - k) \quad (3)$$

$$\Psi_{j,k}(t) = 2^{-j/2}\Psi(2^{-j}t - k) \quad (4)$$

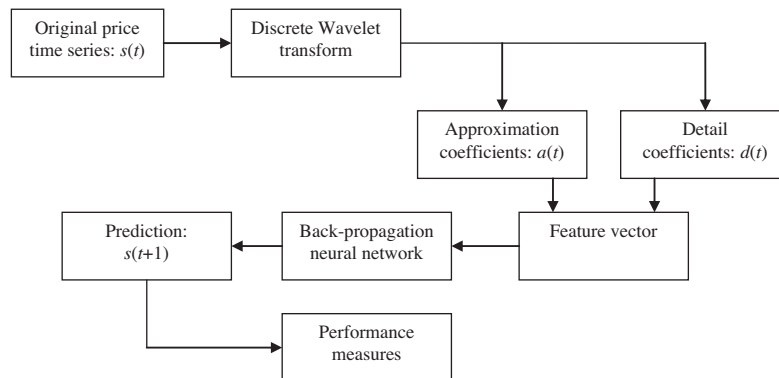


Figure 1 The proposed stock price prediction system.

The two wavelets Φ and Ψ satisfy the following condition:

$$\int \Phi(t) dt = 1 \quad (5)$$

$$\int \Psi(t) dt = 0 \quad (6)$$

As a result, the orthogonal wavelet representation of the signal $s(t)$ is given by

$$s(t) = \sum_k a_{j,k} \Phi_{j,k}(t) + \sum_k d_{j,k} \Psi_{j,k}(t) + \sum_k d_{j-1,k} \Psi_{j-1,k}(t) + \dots + \sum_k d_{1,k} \Psi_{1,k}(t) \quad (7)$$

The decomposition process of the discrete wavelet transform is shown in Fig. 2. For example, the original signal $s(t)$ is decomposed into approximation coefficients $a(t)$ and detail coefficients $d(t)$, by convolving the signal $s(t)$ with a low-pass filter (LP) and a high-pass filter (HP), respectively. The low-pass filtered signal is the input for the next iteration step and so on. The approximation coefficients $a(t)$ contain the general trend (the low-frequency components) of the signal $s(t)$, and the detail coefficients $d(t)$ contain its local variations (the high-frequency components).

In financial time series modeling and forecasting, various types of wavelets can be used, such as the Haar, Mexican Hat, Morlet and Daubechies wavelets (Rao and Bopardikar, 1998; Percival and Walden, 2000). However, the Mexican Hat and the Morlet wavelet are expensive to calculate, and the Haar wavelet is discontinuous and does not approximate continuous signals very well. Moreover, the popular Daubechies wavelet is a compactly supported orthonormal wavelet and provides accurate time series predictions; hence, it is widely used in financial time series forecasting problems (Chang and Fan, 2008; Huang, 2011; Huang and Wu, 2010). In this paper, the daubechies-4 (db4) (Daubechies, 1992) wavelet is applied to decompose the original signal $s(t)$. The level of decomposition is set to two. Thus, the scaling and translation parameter are, respectively, set to two and one. Finally, the Matlab Wavelet Toolbox is employed to perform the DWT on the data.

2.2. Backpropagation neural networks

The multilayer neural networks (NN) that are trained by using the backpropagation (BP) algorithm are the most popular choice in neural network applications in finance. The backpropagation neural networks (BPNN) are feed-forward neural networks with one or more hidden layers, which are capable of approximating any continuous function up to a certain accuracy with only one hidden layer (Cybenko, 1989; Funahashi, 1989). The BPNN consists of three types of layers. The first layer is the input layer and corresponds to the problem's input variables, with one node for each input variable. The second layer is the hidden layer, which is used to capture the non-linear relationships among the variables. The third layer is the output layer, which is used to provide the predicted values. Fig. 3 shows a three-layer BPNN with two neurons (approximation coefficients a and detail coefficients d) in the input layer, four neurons in the hidden layer and one neuron in the output layer. For example, the output

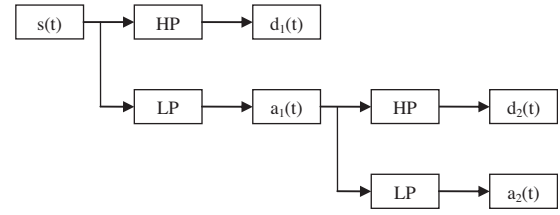


Figure 2 A two-level wavelet decomposition of a signal $s(t)$.

layer has only one neuron, which corresponds to the prediction result. This architecture is adopted in this study.

The relationship between the output $y(t)$ and the input $x(t)$ is given by the following:

$$y(t) = w(0) + \sum_{j=1}^u w(j) \cdot f\left(w(0,j) + \sum_{i=1}^v w(i,j) \cdot x(t)\right) \quad (8)$$

where $w(i,j)$ ($i = 0,1,2,\dots,p$; $j = 1,2,\dots,u$) and $w(j)$ ($j = 0,1,2,\dots,u$) are the connection weights, v is the number of input nodes, u is the number of hidden nodes, and f is a non-linear activation function that enables the system to learn non-linear features. The most widely used activation functions for the output layer are the sigmoid and hyperbolic functions. In this paper, the sigmoid transfer function is employed because it is suitable for fitting our data. It is given by

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

The NN is trained by using the backpropagation (BP) algorithm, and the weights are optimized. The objective function to minimize is the sum of the squares of the differences between the desirable output $y_d(t)$ and the predicted output $y_p(t)$, which is given by

$$E = \frac{1}{2} \sum_{t=1}^n (y_p(t) - y_d(t))^2 \quad (10)$$

The training of the network is performed by the Backpropagation algorithm (Rumelhart et al., 1986), which is trained with the steepest descent algorithm, as follows:

$$\Delta w(k) = -\alpha(k)g(k) + m\Delta w(k-1) \quad (11)$$

where $\Delta w(k)$ is the vector of weight changes, $g(k)$ is the current gradient, $\alpha(k)$ is the learning rate, which determines the length of the weight update, and m is the momentum parameter, which allows escaping from small local minima on the error

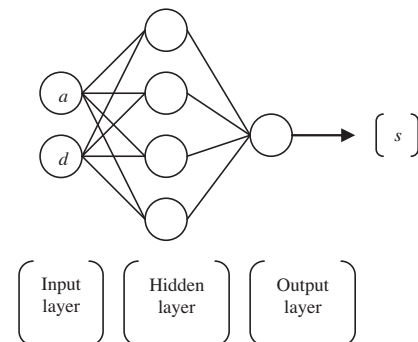


Figure 3 Topology of a one hidden layer BPNN.

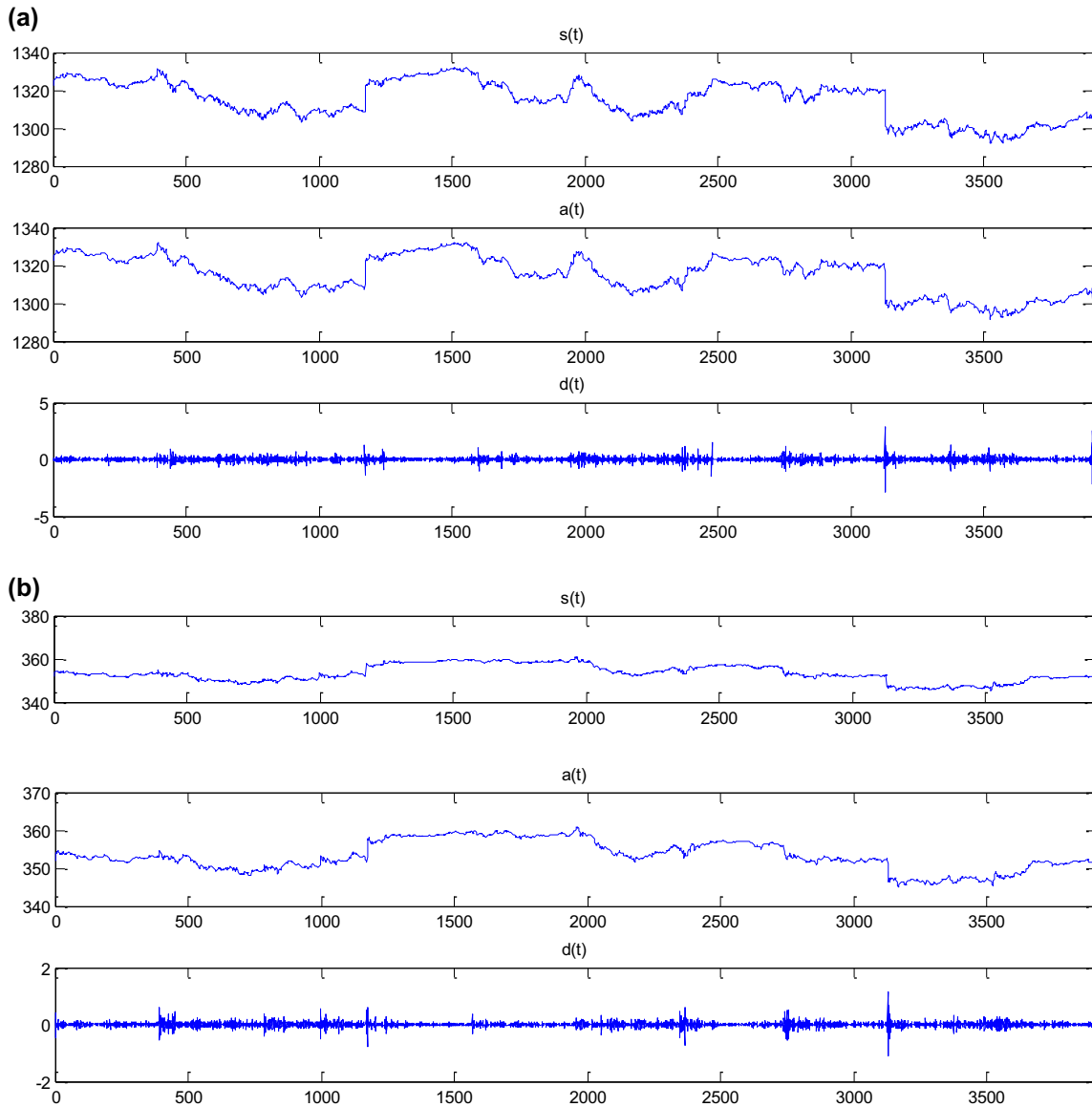


Figure 4 (a) DWT decomposition of the S & P 500. (b) DWT decomposition of Apple. (c) DWT decomposition of Dell. (d) DWT decomposition of Hewlett–Packard. (e) DWT decomposition of IBM. (f) DWT decomposition of Microsoft. (g) DWT decomposition of Oracle.

surface (Ramírez et al., 2003) and avoids having oscillations reduce the sensitivity of the network to fast changes in the error surface (Jang et al., 1997). The learning rate and the momentum parameter are arbitrarily set to 0.01 and 0.9, respectively. The number of epochs that are used to train the BPNN is set to 100. The training of the BPNN will stop when the error E achieves 0.0001 or when the number of epochs reaches 100.

3. Comparison and evaluation criteria

The main reference model in our study is the BPNN, which uses approximation coefficients to predict stock prices. In addition, the auto-regressive moving-average (ARMA) process (Box and Jenkins, 1976) is utilized as a secondary reference

model to predict each stock price. The ARMA process is a popular statistical approach for financial time series forecasting (Denton, 1995; Hann and Steurer, 1996; Taskaya and Casey, 2005). The ARMA(p, q) model represents the future value of a variable as a linear function of past observations and random errors. For example, the ARMA(p, q) process of a time series Z is given by

$$Z_t = c + \sum_{i=1}^p \phi_i Z_{t-i} + a_t + \sum_{j=1}^q \theta_j a_{t-j} \quad (12)$$

where t is the time script, c is a constant term, p is the order of the auto-regressive component, q is the order of the moving average component, ϕ and θ are coefficients to be estimated, and a is a random error that is assumed to be independently and identically distributed with a mean of zero and a constant

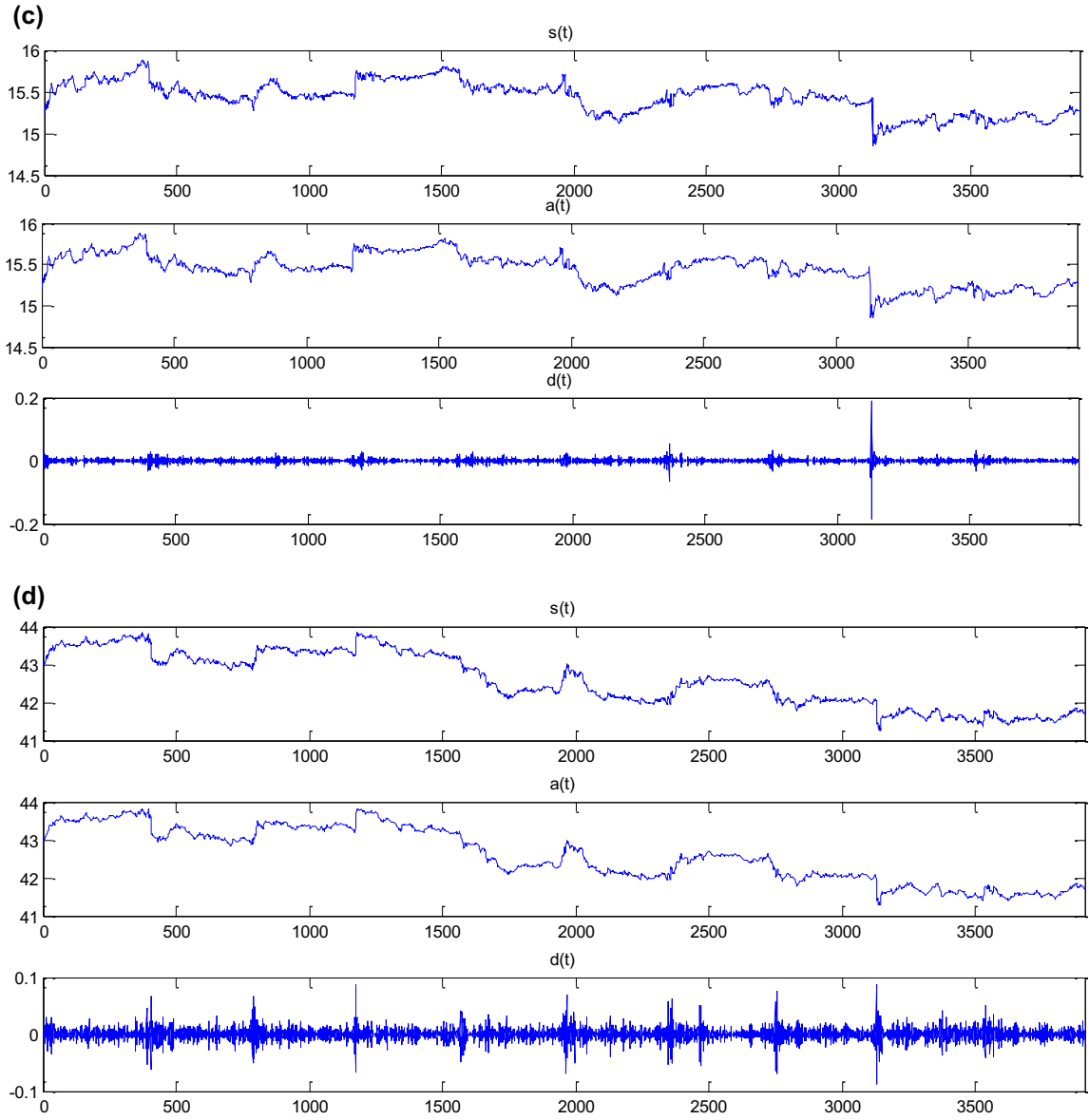


Fig. 4 (continued)

variance of σ_2 . The [Box and Jenkins \(1976\)](#) methodology includes three steps to identify the ARMA parameters p and q . The autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the sample data are used to identify the orders p and q . Finally, Eq. (11) is estimated by maximum likelihood estimation. More details about ARMA modeling can be found in [Hamilton \(1994\)](#). In addition, the random walk process (RW) is also used for comparison purposes. It is defined as follows:

$$Z_t = c + a_t \quad (13)$$

where the parameter c and the variable a have been defined previously. Finally, to evaluate the performance of the BPNN, ARMA model and the random walk process, three evaluation criteria are used as accuracy measures, namely, the mean absolute error (MAE), the root mean-square error (RMSE) and the mean absolute deviation (MAD). These evaluation criteria are calculated as follows:

$$\text{MAE} = \frac{1}{N} \sum_{t=1}^N |s(t) - p(t)| \quad (14)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (s(t) - p(t))^2} \quad (15)$$

$$\text{MAD} = \frac{1}{N} \sum_{t=1}^N |s(t) - \bar{p}(t)| \quad (16)$$

where $s(t)$, $p(t)$ and $\bar{p}(t)$ are, respectively, the true signal, the forecasted signal and the average of the forecasted signal over the testing (out-of-sample) period $t = 1$ to N .

The smaller the values of these performance measures, the closer are the forecasted signal values to the true signal values. In other words, the lower the evaluation criteria are, the better the performance in forecasting.

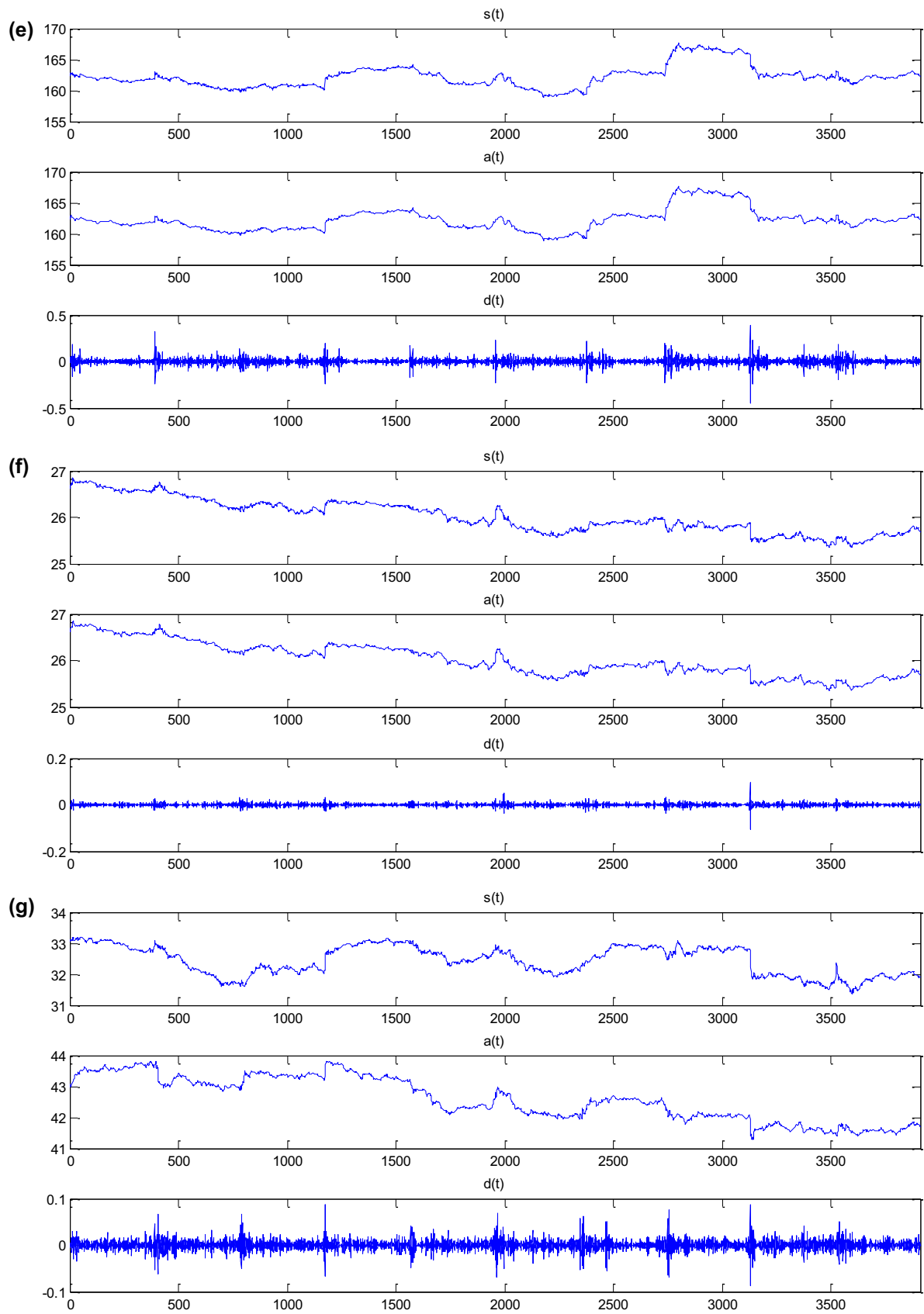


Fig. 4 (continued)

4. Data and results

To evaluate the performance of our proposed stock price prediction system, the S & P 500 price index and six stock prices are used. The stocks are Apple, Dell, Hewlett–Packard, IBM, Microsoft and Oracle. The data were downloaded from the Yahoo finance website. All of the data are minute-in-day closing prices for the period from February 28th 2011 to March 11th 2011. There are 391 min of quotations per day. Accordingly, there are a total of 3910 data points for each dataset. The first 3128 data points (80% of the total sample points) are used as the training sample, while the remaining 782 data points (20% of the total sample points) are used as the testing sample. For each dataset, the price series $s(t)$ and their approximation $a(t)$ and detail $d(t)$ coefficients are shown in Fig. 4a–g.

The forecasting results of the proposed model (BPNN + AC + DC) are shown in Table 1. For all of the times series, we found that when the orders p and q are set to one, the ARMA better fits the data. Clearly, it is shown that the performance measures (MAE, RMSE, MAD) obtained with backpropagation neural networks (BPNN) using both

approximation (AC) and detail coefficients (DC) are smaller than those obtained with the standard approach, which is based only on approximation coefficients (BPNN + AC). This effect is very pronounced on the S & P 500 and Hewlett–Packard. For example, using the S & P 500 dataset, the MAE, RMSE and MAD obtained with the standard approach, which is based on BPNN and approximation coefficients (AC), are, respectively, 8.2984, 8.3818 and 0.8872. In contrast, the values obtained with our approach (BPNN + AC + DC) are, respectively, 0.000327, 0.000329 and 0.000019. This result can be explained by the fact that the S & P 500 is a market index that is composed of many less volatile companies from different sectors in the US economy. In other words, the S & P 500 is less volatile than the individual technology companies used in our studies. Hence, both the standard approach (BPNN + AC) and our model (BPNN + AC + DC) were able to predict the S & P 500 with very small errors. For Hewlett–Packard, the obtained MAE, RMSE and MAD with the standard approach (versus our model) are, respectively, 2.6099 (0.3186), 2.6329 (0.3407) and 0.2834 (0.0988).

Similar results are obtained with the other companies. For the prediction of Apple, the obtained MAE, RMSE and MAD with the standard approach (versus our model) are, respectively, 4.2955 (4.2681), 5.3018 (4.7162) and 4.7957 (1.8152). For the prediction of Dell, the obtained MAE, RMSE and MAD with standard BPNN + AC (versus our model BPNN + AC + DC) are, respectively, 0.1293 (0.0565), 0.1425 (0.0751) and 0.0511 (0.0387). For the prediction of IBM, the obtained MAE, RMSE and MAD with standard BPNN + AC (versus BPNN + AC + DC) are, respectively, 3.3756 (2.1185), 3.4145 (2.1275) and 0.3826 (0.1340). For the prediction of Microsoft, the obtained MAE, RMSE and MAD with standard BPNN + AC (BPNN + AC + DC) are, respectively, 0.6987 (0.2785), 0.712 (0.292) and 0.1101 (0.0693). Finally, for the prediction of Oracle, the obtained MAE, RMSE and MAD with standard BPNN + AC (BPNN + AC + DC) are, respectively, 1.7091 (1.1365), 1.7196 (1.1473) and 0.1521 (0.1258).

In summary, for all of the datasets, the deviations between the actual and predicted values are smaller when the proposed model (BPNN + AC + DC) is applied to both the stock market and the individual shares. Thus, the detail coefficients (DC) help to improve the forecasting accuracy. Indeed, in the wavelet space, the detail coefficients are features that reveal the properties that cannot be detected by the approximation coefficients (AC). Consequently, detail coefficients (DC) act as a complementary part of the approximation coefficients (low-frequency components). As a result, better forecasting performance is achieved with our model (BPNN + AC + DC), which is based on both detail and approximation coefficients as predictive patterns. Thus, our model outperforms the standard model (BPNN + AC), which uses approximation coefficients (AC) only as predictive inputs to be given to the BPNN.

Finally, both the standard approach, which is based on BPNN with low-frequency components and our model (which extends the latter by integrating high-frequency components) outperforms the statistical approach, which is based on the conventional ARMA model and the random walk process (RW) (see Table 1).

Although previous studies have reported that ARMA models outperform artificial neural networks in terms of forecast-

Table 1 Simulation results.

	MAE	RMSE	MAD
<i>S & P 500</i>			
RW	2.7949	3.4525	1.8645
ARMA	31.9292	155.7031	35.6351
BPNN + AC	8.2984	8.3818	0.8872
BPNN + AC + DC	0.000327	0.000329	0.000019
<i>Apple</i>			
RW	5.9097	6.3141	4.9454
ARMA	7.8568	41.7257	9.5801
BPNN + AC	4.2955	5.3018	4.7957
BPNN + AC + DC	4.2681	4.7162	1.8152
<i>Dell</i>			
RW	0.3443	0.3517	1.5099
ARMA	0.4420	1.8414	0.4177
BPNN + AC	0.1293	0.1425	0.0511
BPNN + AC + DC	0.0565	0.0751	0.0387
<i>Hewlett–Packard</i>			
RW	0.3481	0.3488	0.0996
ARMA	0.7900	4.9438	1.1415
BPNN + AC	2.6099	2.6329	0.2834
BPNN + AC + DC	0.3186	0.3407	0.0988
<i>IBM</i>			
RW	3.704	4.842	1.8474
ARMA	4.0278	19.3699	4.4304
BPNN + AC	3.3756	3.4145	0.3826
BPNN + AC + DC	2.1185	2.1275	0.1340
<i>MICROSOFT</i>			
RW	0.5655	0.5728	0.2060
ARMA	0.5033	3.0390	0.7008
BPNN + AC	0.6987	0.712	0.1101
BPNN + AC + DC	0.2785	0.292	0.0693
<i>ORACLE</i>			
RW	7.435	7.608	2.629
ARMA	1.0413	3.8751	0.8726
BPNN + AC	1.7091	1.7196	0.1521
BPNN + AC + DC	1.1365	1.1473	0.1258

ing accuracy (Denton, 1995; Hann and Steurer, 1996; Taskaya and Casey, 2005), our finding is in accordance with Zhang (2003) and Khashei and Bijari (2012), who found that BPNN outperforms the ARMA process. This result can be explained as follows. In the one hand, the popular ARMA model that is widely applied in the financial industry assumes a linear relationship between current and past values of a time series as well as with white noise. Therefore, it fits linear relations better than nonlinear ones. On the other hand, BPNNs are nonlinear models that use information from the wavelet domain to predict future stock prices. The wavelet domain information contains both long and short patterns to characterize a financial time series. The bottom line is that our model BPNN + AC + DC outperforms the standard model BPNN + AC, the conventional statistical ARMA model and the random walk process.

5. Conclusions

In investment decision making, stock price prediction is an important activity for both financial firms and private investors. The first step of a stock price prediction model consists of extracting features. In recent years, the discrete wavelet transform was largely used for the extraction of information contained in stock price time series. In particular, low-frequency approximation coefficients were used to predict future stock prices. However, approximation components characterize only the coarse structure of data to identify the long-term trends. To account for local information such as discontinuities, ruptures and singularities in the original data, detail coefficients should also be extracted to serve as additional inputs to predicting stock prices.

This paper presented a forecasting model that integrates discrete wavelet transforms and backpropagation neural networks for financial time series. The presented model first uses wavelet transforms to decompose the financial time series data. Then, the obtained approximation and detail components after decomposition of the original time series are used as input variables to forecast future stock prices.

Our simulation results showed that the low-frequency components coupled with high-frequency components resulted in higher accuracy compared to a conventional model that uses only low-frequency components to predict future stock prices. In addition, our model outperformed both the well-known statistical ARMA model and the random walk process. The presented model was shown to be effective in financial forecasting. It can also be implemented for real-time prediction because the data processing time is less than one minute.

For future work, we aim to examine the effect of both the wavelet choice and the level of decomposition on the accuracy of different machine learning techniques, such as the back-propagation neural networks used in our study, recurrent neural networks and support vector machines, to name a few.

Acknowledgment

The author gratefully thanks Stephane Gagnon from the University of Quebec at Outaouais (Canada) for providing the data for this work.

References

- Atsalakis, G.S., Valavanis, K.P., 2009. Surveying stock market forecasting techniques – Part II: soft computing methods. *Expert Syst. Appl.* 36, 5932–5941.
- Bahrammirzaee, A., 2010. A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. *Neural Comput. Appl.* 19, 1165–1195.
- Box, G.E.P., Jenkins, G., 1976. *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco.
- Chang, P.C., Fan, C.Y., 2008. A hybrid system integrating a wavelet and TSK fuzzy rules for stock price forecasting. *IEEE Trans. Syst., Man, Cybernet.* – Part C: Appl. Rev. 38, 802–815.
- Cybenko, G., 1989. Approximations by superpositions of a sigmoidal function. *Math. Control Signals Systems* 2, 303–314.
- Daubechies, I., 1990. The wavelet transform: time–frequency localization and signal analysis. *IEEE Trans. Inf. Theory* 36, 961–1005.
- Daubechies, I., 1992. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, Pennsylvania, USA.
- De, I., Sil, J., 2012. Entropy based fuzzy classification of images on quality assessment. *J. King Saud Univ. – Comput. Inf. Sci.* 24, 165–173.
- Denton, J.W., 1995. How good are neural networks for causal forecasting? *J Business Forecast.* 14, 17–20.
- Engle, R., Russell, J., 2010. Analysis of high frequency financial data. *Handbook of Financial Econometrics, volume 1: Tools and Techniques*, 383–426.
- Fama, E.F., 1965. Random walks in stock market prices. *Financ. Anal. J.* 21, 55–59.
- Funahashi, K.-I., 1989. On the approximate realization of continuous mappings by neural networks. *Neural Netw.* 2, 183–192.
- Hamilton, J.D., 1994. *Time Series Analysis*. Princeton University Press.
- Hann, T.H., Steurer, E., 1996. Much ado about nothing? Exchange rate forecasting: neural networks vs. linear models using monthly and weekly data. *Neurocomputing* 10, 323–339.
- Hsieh, T.-J., Hsiao, H.-F., Yeh, W.-C., 2011. Forecasting stock markets using wavelet transforms and recurrent neural networks: an integrated system based on artificial bee colony algorithm. *Appl. Soft Comput.* 11, 2510–2525.
- Huang, S.-C., 2011. Forecasting stock indices with wavelet domain kernel partial least square regressions. *Appl. Soft Comput.* 11, 5433–5443.
- Huang, S.C., Wu, T.K., 2008. Combining wavelet-based feature extractions with relevance vector machines for stock index forecasting. *Expert Syst.* 25, 133–149.
- Huang, S.-C., Wu, T.-K., 2010. Integrating recurrent SOM with wavelet-based kernel partial least square regressions for financial forecasting. *Expert Syst. Appl.* 37, 5698–5705.
- Jang, J.-S.R., Sun, C.-T., Mizutani, E., 1997. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. PrenticeHall, NJ.
- Kao, L.-J., Chiu, C.-C., Lu, C.-J., Chang, C.-H., 2013. A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting. *Decis. Support Syst.* 54, 1228–1244.
- Khashei, M., Bijari, M., 2012. A new class of hybrid models for time series forecasting. *Expert Syst. Appl.* 39, 4344–4357.
- Lahmiri, S., 2013. Forecasting direction of the S&P500 movement using wavelet transform and support vector machines. *Int. J. Strateg. Decis. Sci.* 4, 78–88.
- Li, J., Shi, Z., Li, X., 2006. Genetic programming with wavelet-based indicators for financial forecasting. *Trans. Inst. Meas. Control* 28, 285–297.

- Mallat, S.G., 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 11, 674–693.
- Percival, D.B., Walden, A.T., 2000. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge, UK.
- Ramírez, E., Castillo, O., Soria, J., 2003. Hybrid system for cardiac arrhythmia classification with fuzzy k-nearest neighbors and multi layer perceptrons combined by a fuzzy inference system. *The International Joint Conference on Neural Networks (IJCNN)*, 18–23 July, Barcelona, Spain.
- Rao, R.M., Bopardikar, A.S., 1998. *Wavelet Transforms: Introduction to Theory and Applications*. Addison Wesley, Boston.
- Rikli, N.-E., 2012. Self-similarity and stationarity of increments in VBR video. *J King Saud Univ. – Comput. Inf. Sci.* 24, 7–16.
- Rout, M., et al., 2014. Forecasting of currency exchange rates using an adaptive ARMA model with differential evolution based training. *J. King Saud Univ. – Comput. Inf. Sci.* 26, 7–18.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature* 323, 533–536.
- Taskaya, T., Casey, M.C., 2005. A comparative study of autoregressive neural network hybrids. *Neural Netw.* 18, 781–789.
- Wang, J.-Z., Wang, J.-J., Zhang, Z.-G., Guo, S.-P., 2011. Forecasting stock indices with back propagation neural network. *Expert Syst. Appl.* 38, 14346–14355.
- Zhang, G., Patuwo, B.E., Hu, M.Y., 1998. Forecasting with artificial neural networks: the state of the art. *Int. J. Forecast.* 14, 35–62.
- Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 50, 159–175.