Research Article

Multiband Prediction Model for Financial Time Series with Multivariate Empirical Mode Decomposition

Md. Rabiul Islam,[1] Md. Rashed-Al-Mahfuz,[1] Shamim Ahmad,[2] and Md. Khademul Islam Molla[2, 3]

[1] *Department of Computer Science and Engineering, Pabna Science and Technology University, Pabna 6600, Bangladesh*

[2] *Department of Computer Science and Engineering, University of Rajshahi, Rajshahi 6205, Bangladesh*

[3] *Geophysical Sciences, University of Alberta, Edmonton, AB, Canada T6G 2G7*

Correspondence should be addressed to Md. Khademul Islam Molla, mdkhadem@ualberta.ca

This paper presents a subband approach to financial time series prediction. Multivariate empirical mode decomposition (MEMD) is employed here for multiband representation of multichannel financial time series together. Autoregressive moving average (ARMA) model is used in prediction of individual subband of any time series data. Then all the predicted subband signals are summed up to obtain the overall prediction. The ARMA model works better for stationary signal. With multiband representation, each subband becomes a band-limited (narrow band) signal and hence better prediction is achieved. The performance of the proposed MEMD-ARMA model is compared with classical EMD, discrete wavelet transform (DWT), and with full band ARMA model in terms of signal-to-noise ratio (SNR) and mean square error (MSE) between the original and predicted time series. The simulation results show that the MEMD-ARMA-based method performs better than the other methods.

1. Introduction

Time series prediction is used in a wide range of applications, including the business and economic environment, currency forecasting, stock exchange operations in order to develop proper strategies and avoid the risk of potential losses. One step ahead prediction is used in a number of applications such as stock exchange, traffic, or water stream flow forecast. One aim is to estimate a time series value $s(t)$ known past values $s(t-1), s(t-2), \ldots, s(t-k)$. Most of the systems are considered linear relationships among the variables to model this type of time series known as autoregressive (AR) models based on the Box and Jenkins method

[1, 2]. Recently, stock markets forecasting is interesting field for the investors and it has got very high concentration in financial time series analysis and financial researchers. To invest and take proper decision, more precise forecasting of financial environments is an important issue. In conventional econometric models, the variance of the disturbance is assumed to be constant. However, financial time series data exhibit periods of unusual large volatility (nonstationary and noisy) and considered as more challenging task for analysis [3, 4].

There are many forecasting models that have been developed in predicting financial time series. Among them, most popular technique of linear analysis is perhaps autoregressive moving average (ARMA) method [2] for stationary time series. Usually, time series data are nonstationary and appear with noises that reduce the prediction accuracy of ARMA approach. The efficiency of ARMA model is improved here by using multiband prediction model. The multiband forecasting performance of financial data is better than full-band approach. The multiband decomposition is usually performed by employing Fourier transform (FT) [5–7] or wavelet transform (WT) [8–10] based filter bank. Fourier transform is not suitable to adapt with the non-stationary signal, whereas wavelet transform is relatively more data adaptive decomposition [11]. Both of the transformations use priori bases for signal decomposition. The FT assumes any signal as a sum of sinusoids and WT adapts its basis with the signal by changing amplitude and scale of the basis. In Fourier spectra, the existence of energy at a frequency means a component of sinusoidal wave of that frequency persisted through the time span of the data and it also includes the harmonic distortion with a nonharmonic signal [12]. Although discrete wavelet leads to nonredundant signal decomposition [13], the wavelet spectra gives a smeared average frequency range over which the main signal energy resides [12]. Hence, both transformations have some limitations in multiband decomposition of nonstationary signal.

Sampling is an important factor for modeling and analysis of financial data because the analysis is performed in a discrete way. Any financial signal $x(t)$ at time instant t can be defined as $x(t) \Rightarrow x(nT)$, where n is a nonnegative number representing the index of the discrete sample and T is the sampling period. For uniform sampling rate, T can be omitted and then $x(t) \Rightarrow x(n)$, with n representing the discrete sample index. The data used in this study are collected with uniform sampling over total period of acquisition. The nonuniform sampling/discretization is always better to adapt with the signal variation. The basic concept of nonuniform sampling is to increase the sampling rate to acquire the signal with higher variation and the reverse when the lower variation of signal is observed in acquisition. The current research trend is to apply the nonuniform discretization to increase the efficiencies of different systems [14, 15]. The stabilization of linear time-invariant dynamic system is achieved by applying multirate discretization combined with fractional-order hold [14]. The stabilization of such system is not guaranteed without using multirate sampling. The controllability and observability are investigated for Caputo fractional differential linear system of any real order α in [15]. The developed model also supports those properties under nonuniform sampling. The authors have also proved that the choice of appropriate sampling instance is not restrictive as a result of the properties of associate Chebyshev's systems. Since we are using the secondary data, the sampling rate is kept as it is although the nonuniform sampling has many advantages.

The proposed method focuses on the study of multiband ARMA model in which the multiband representation is implemented by using empirical mode decomposition (EMD), a fully data adaptive decomposition technique. The EMD is a recently developed method for multiband representation of nonlinear and nonstationary time series [12]. It is also considered as a dyadic filter bank and the IMF components are all normally distributed [16]. The

specialty of this decomposition is that its bases are data adaptive and it is fully reversible, that is, the original data is reconstructed simply by summing up the bases with a negligible error. Although EMD-based decomposition method is fully a data adaptive method, its mode mixing is evident between intrinsic mode functions (IMFs) or a single mode is "leaked" into two IMFs. The extensions of EMD to bivariate and trivariate EMD cases reported in [17, 18] are limited to handle only two and three variables, respectively. To analyze any multivariate data it is necessary to make selection of variables to fit the use of BEMD and TEMD for more than two and three variables, respectively. The performance is very much dependent on such selection of variables for processing at a time. The MEMD algorithm is the first generic extension of standard EMD for multivariate data and its important step to generate multiple n-dimensional envelopes by taking signal projections along different directions in n-dimensional spaces. It is noted that MEMD reduces the mode mixing problem of the basis functions obtained from multivariate data. In this paper, multiband ARMA model with MEMD is introduced in prediction of financial time series. Recently, financial time series forecasting using classical EMD is introduced in [19] for only single-channel data. Any model to handle multichannel data is not mentioned. There is no guarantee that same number of IMFs will be produced for different data channels using classical EMD (which leads to different prediction model for different channels), whereas the use of MEMD ensures this mater.

The financial data is decomposed into multiband using discrete wavelet transformation (DWT), EMD, and MEMD. The ARMA parameters are computed for individual band and the overall prediction is obtained by summing up the predicted signal of all bands. The prediction error measured in terms of signal-to-noise ratio (SNR), mean square error (MSE), and the rose histogram of the absolute differences of the computed and actual time series. The performance of the proposed method is compared with other related techniques. Regarding the arrangement of this paper, the multiband decomposition techniques using DWT, EMD, and MEMD are described in Section 2, the prediction model introduced in Section 3, several performance criteria are described in Section 4, the experimental results are illustrated Section 5, and finally discussion and concluding remarks are presented in Section 6.

2. Multiband Decomposition

2.1. Discrete Wavelet Transformation

Discrete wavelet transform (DWT) is also widely used to perform the multiband representation of financial data. Reconstruction of the original signal from the details obtained by wavelet filterbank is more convenient than Fourier-based method. It is also possible to design perfect reconstruction filterbank in Fourier domain, whereas wavelet filterbank automatically satisfies such requirement. It is well known in the subband filtering community that if the same FIR filters are used for reconstruction and decomposition, then symmetry and exact reconstruction are incompatible [20]. Therefore, biorthogonal wavelet which incorporates separate bases for decomposition and reconstruction is used here. Two bases of a biorthogonal are shown in Figure 1.

One wavelet $\widetilde{\psi}$ is used in the analysis, and the coefficients of a signal $s(t)$ are given as

$$\widetilde{c}_{j,k} = \sum_t s(t)\widetilde{\psi}_{j,k}(t), \tag{2.1}$$
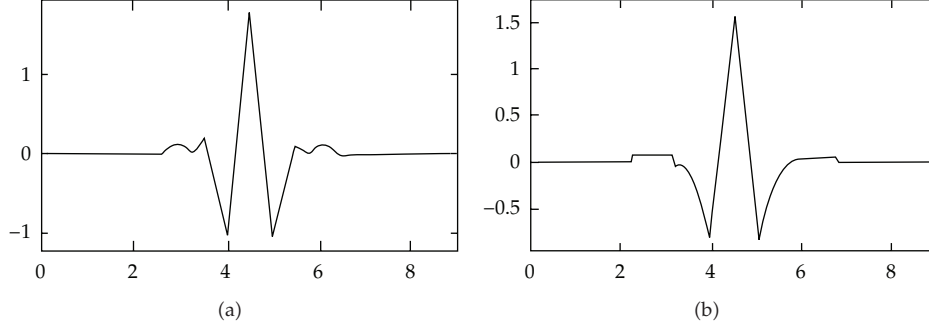
Figure 1: Biorthogonal wavelet (db8) for decomposition (a) and reconstruction (b).

where j and k are the scaling and shifting factors, respectively. The other wavelet ψ is used for reconstruction as

$$s = \sum_{j,k} \widetilde{c}_{j,k} \psi_{j,k}. \tag{2.2}$$

Furthermore, the two wavelets are related by duality in the following sense:

$$\sum_t \widetilde{\psi}_{j,k}(t) \psi_{j',k'}(t) = 0 \quad \text{for } j \neq j' \text{ or } k \neq k'. \tag{2.3}$$

To aid the development, the so-called scaling function $\widetilde{\phi}$ is used to approximate the signal up to a particular level of detail. A family of scaling functions for decomposition can be represented via shifts and power-of-two stretches as

$$\forall j, k, j \wedge k \in Z : \widetilde{\phi}_{j,k}(t) = 2^{-(j/2)} \widetilde{\phi}\left(2^{-j}t - k\right). \tag{2.4}$$

The dual scaling function ϕ is used for reconstruction and the relation between $\widetilde{\phi}$ and ϕ can be represented as

$$\sum_t \widetilde{\phi}_{0,k}(t) \phi_{0,k'}(t) = 0 \quad \text{for } k \neq k'. \tag{2.5}$$

Hence the 4-touple $(\widetilde{\psi}, \widetilde{\phi}, \psi, \phi)$ can be considered as a compactly supported biorthogonal wavelets with scaling functions [20]. It becomes apparent, as in [21], that the useful properties for analysis can be concentrated in the $\widetilde{\psi}$ function, whereas the interesting properties for reconstruction are assigned to the ψ function. $\widetilde{\psi}$ and ψ can have very different regularity properties, ψ being more regular than $\widetilde{\psi}$. The $\widetilde{\psi}$, ψ, $\widetilde{\phi}$, and ϕ functions are zero outside the segment. The decomposition of the financial data illustrated in Figure 2(a) into four sub-bands and original financial data, its reconstructed data, and the reconstruction error are illustrated by Figure 2(b). The signal-to-noise ratio (SNR) of the reconstructed signal (with respect to original one) is 305 dB.
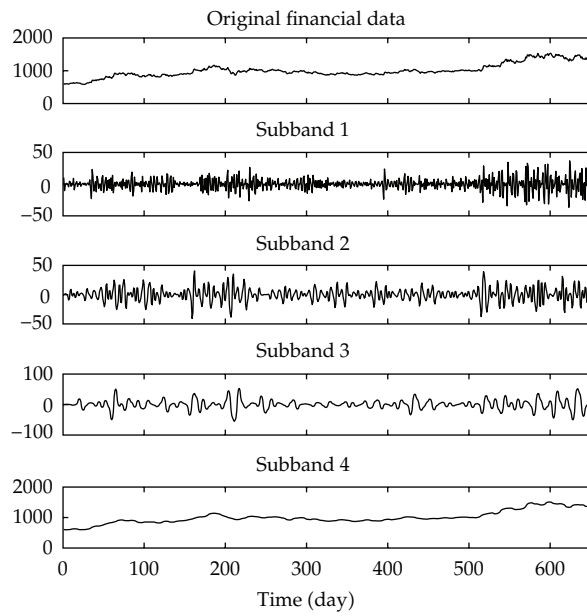
(a)
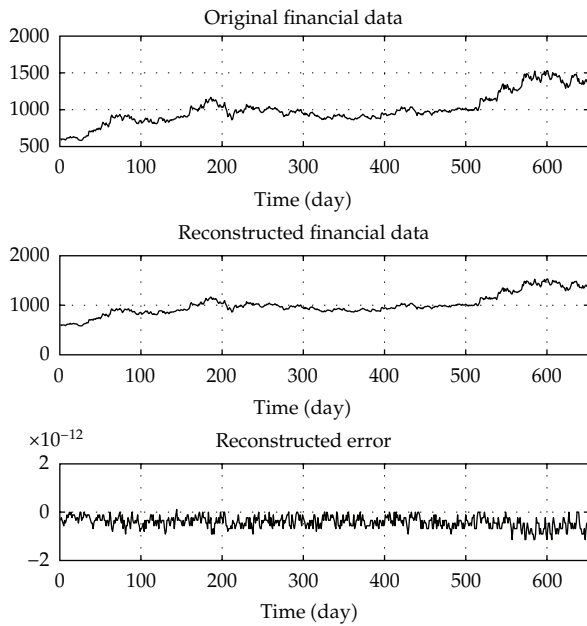
(b)

Figure 2: (a) Wavelet decomposition of financial data into four subbands. (b) Original financial data for channel 1 (top), reconstructed data by simply adding the subbands (middle) and reconstruction error (bottom). The SNR of the reconstructed signal is 305 dB.

2.2. Classical EMD

The key benefit of using EMD is that it is an automatic decomposition and fully data adaptive. The principle of the EMD technique is to decompose a signal $s(t)$ into a finite set of band-limited functions $C_m(t)$ called intrinsic mode functions (IMFs). Each IMF is considered as a AM-FM oscillatory component satisfying two basic conditions: (i) in the whole data set, the number of extrema and the number of zero crossings must be the same or differ at most by one, (ii) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. The first condition is similar to the narrowband requirement for a stationary Gaussian process and the second condition is a local requirement induced from the global one and is necessary to ensure that the instantaneous frequency will not have redundant fluctuations as induced by asymmetric waveforms. However, a special "sifting" process is employed to extract all of the IMFs and this sifting process is described as following. Firstly, the upper and lower envelopes of the signals $s(t)$, as well as their mean value $\mu_1(t)$, are calculated, respectively. The first step of the sifting process is to calculate the difference $h_1(t) = s(t) - \mu_1(t)$. However, $h_1(t)$ rarely satisfies the two IMF properties and is not taken as the first IMF of the signals straightway. Therefore, the sifting usually has to be implemented for more times, where the "difference" obtained in the previous sifting is taken as "signals" in present sifting. If after $(d + 1)$th sifting, corresponding difference, $h_{1d}(t) = h_{1(d-1)}(t) - \mu_{1d}(t)$, satisfies the IMF properties, then it can be taken as the first IMF component, denoted by $C_1(t)$, that is, $C_1(t) = h_{1d}(t)$. In practice, to determine whether or not $h_{1d}(t)$ well-satisfies the IMF properties, we usually use so-called standard deviation (δ) criterion, that is, to check if the following inequality holds [12]:

$$\delta_d = \frac{\sum_{t=1}^{T} \left| h_{1(d-1)}(t) - h_d(t) \right|^2}{\sum_{t=1}^{T} \left| h_{1(d-1)}(t) \right|^2} \leq \eta, \tag{2.6}$$

where T is the frame length and $0.2 \leq \eta \leq 0.3$. Next, taking rest data $r_1(t) = s(t) - C_1(t)$ as "new" signals and implementing the sifting process on it, we can obtain the second IMF $C_2(t)$. This procedure should be repeatedly used for M (total number of IMF components) times until the last residue $r_M(t)$ becomes a monotonic function. At the end of the decomposition the signal $s(t)$ is represented as

$$s(t) = \sum_{m=1}^{M} C_m(t) + r_M(t), \tag{2.7}$$

where $C_1(t), C_2(t), \ldots, C_M(t)$ are all of the IMFs included in the signals and $r_M(t)$ is the final residue.

Another way to explain how EMD works is that it extracts out the highest frequency oscillation that remains in the signal. Thus locally, each IMF contains lower frequency oscillation than the one extracted just before. Being data adaptive, the basis usually offers a physically meaningful representation of the underlying processes. There is no need of considering the signal as stack of harmonics and, therefore, EMD is ideal for analyzing nonstationary and nonlinear data. Each IMF is considered as a monocomponent contribution such that the derivation of instantaneous amplitude and frequency provides a physical significance. The advantage of this time-space filtering is that the resulting band passed

signals preserve the full nonstationary property in physical space. This filtering method is intuitive and direct its basis is a posteriori and data adaptive. The completeness of the decomposition is given by (2.7). The original signal can easily be reconstructed by simply adding the bases (including the residue) with negligible error term. The EMD of a financial data is shown in Figure 3(a). As an example consider the 6th IMF contains a part of the financial data of lower frequency than the signal contained by the 5th IMF. The original signal and reconstructed signal and error (of the order $10 \times e^{-13}$) are shown in Figure 3(b). The signal-to-noise ratio (SNR) of the reconstructed signal (with respect to original one) is 318 dB which is the very negligible error. It is also observed that higher-order IMFs contain lower frequency oscillations than that of lower-order IMFs.

2.3. Multivariate EMD

The Empirical Mode Decomposition (EMD) is a signal analysis technique that decomposes any signal into basis functions modulated in both amplitude and frequency by extracting all of the oscillatory modes embedded in the signal [12]. The decomposition is a data adaptive and signal-dependent decomposition and it does not require any condition about the stationarity and linearity of the signal. The bases extracted by EMD are named intrinsic mode functions (IMFs). Each IMF is symmetric and it is assumed to yield a meaningful local frequency traces. Different IMFs do not exhibit the same frequency at the same time.

    The multivariate EMD (MEMD) is more generalized extension of the EMD suitable for dealing with direct processing of multivariate data including synthetic sinusoidal signals for its real-world applications [22]. Standard EMD revealed that IMFs tend to mimic a filter bank-like decomposition, similar to wavelet decompositions [16, 23]. Multivariate EMD not only exposes filter bank structure but also ensures the overlapping of the frequency responses of the filters associated with the IMFs of the same index from multiple channels. To extend general idea of multivariate signals for MEMD, input data are straightforwardly processed in n-dimensional spaces to generate multiple n-dimensional envelopes by taking signal projections along different directions in n-dimension spaces. The calculation of the local mean can be considered an approximation of the integral of all the envelopes along multiple directions in an n-dimensions space [22, 24]. This step is complex to perform due to the lack of formal definition of maxima and minima in n-dimensional domains in general EMD. The sampling based on low discrepancy Hammersley sequence is used to generate projections of input signal in [22]. Once the projections along different directions in multidimensional spaces are obtained, their extrema are interpolated via cubic spline interpolation to obtain multiple-signal envelopes. Thus obtained envelopes are then averaged to obtain the local mean of the multivariate signal. The following algorithm proposed in [25] is employed here to decompose signal $s(t)$ into a set of IMF components.

(1) Generate the point-set-based on the Hammersley sequence for sampling on an $(n - 1)$-sphere [22].

(2) Calculate a projection, denoted by $\{p^{\theta_k}(t)\}_{t=1}^{T}$, of the input signal $\{s(t)\}_{t=1}^{T}\}$ along the direction vector X^{θ_k}, for all k (the whole set of direction vectors), giving $\{p^{\theta_k}(t)\}_{k=1}^{K}$ as the set of projections.

(3) Find the time instants $\{t_i^{\theta_k}\}_{k=1}^{K}$ corresponding to the maxima of the set of projected signals $\{p^{\theta_k}(t)\}_{k=1}^{K}$.
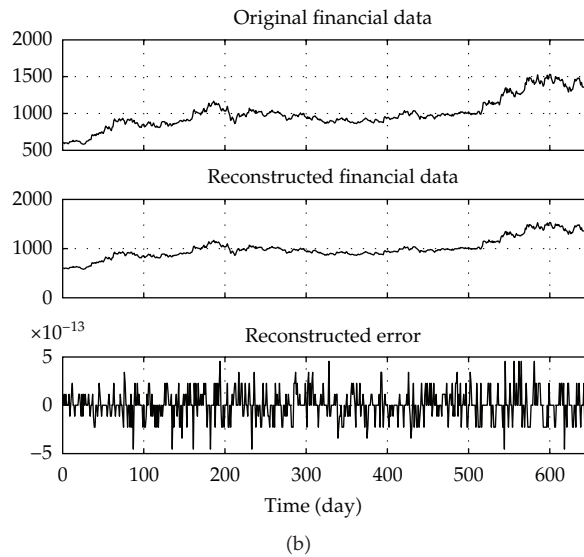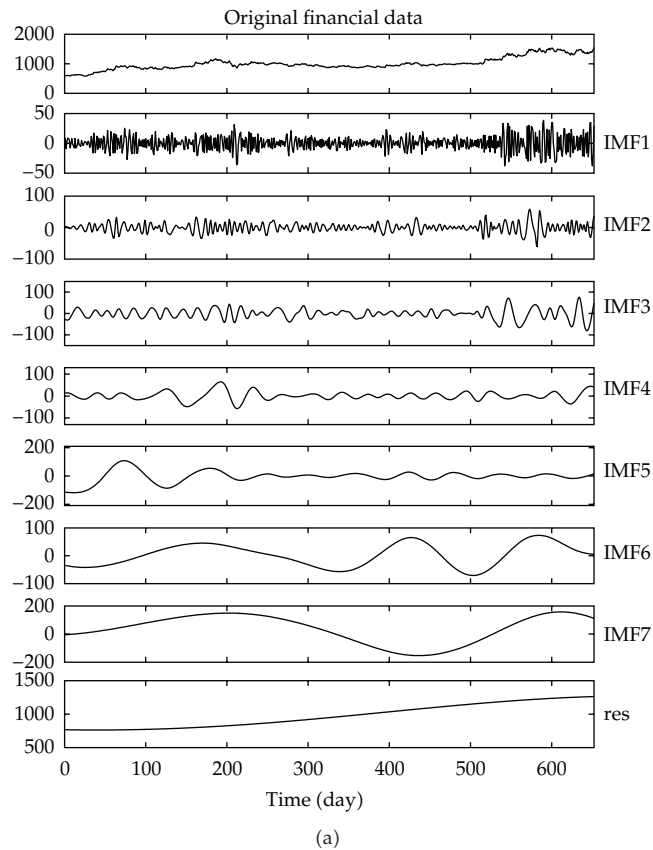
(a)

(b)

Figure 3: (a) EMD (different IMFs) of financial data. (b) Original financial data for channel 1 (top), reconstructed data by simply adding the IMFs (middle), and reconstruction error (bottom). The SNR of the reconstructed signal is 318 dB.

(4) Interpolate $[t_i^{\theta_k}, s(t_i^{\theta_k})]$, for all values of k, to obtain multivariate envelope curves $\{e^{\theta_k}(t)\}_{k=1}^{K}$.

(5) For a set of K direction vectors, calculate the mean $\mu(t)$ of the envelope curves as

$$\mu(t) = \frac{1}{k} \sum_{k=1}^{K} e^{\theta_k}(t). \tag{2.8}$$

(6) Extract the "detail" $d(t)$ using $d(t) = X(t) - \mu(t)$. If the "detail" $d(t)$ fulfills the stoppage criterion for a multivariate IMF, apply the above procedure to $X(t) - d(t)$, otherwise apply it to $d(t)$.

Consider a sequence of N-dimensional vectors $\{s(t)\}_{t=1}^{T} = \{s_1(t), s_2(t), \dots, s_N(t)\}$ representing a multivariate signal with N components, and $X^{\theta_k} = \{x_1^k, x_2^k, \dots, x_N^k\}$ denoting a set of direction vectors along the directions given by angles $\theta^k = \{\theta_1^k, \theta_2^k, \dots, \theta_{(N-1)}^k\}$ on an $(n-1)$-sphere.

Once the first IMF is extracted, it is subtracted from the input signal and the same process is applied to the resulting signal yielding the second IMF and so on. In the multivariate case, the residue corresponds to a signal whose projections do not contain enough extrema to form a meaningful multivariate envelope. The stopping criterion for MEMD of IMFs is similar to standard EMD [26], the difference being that the condition for equality of the number of extrema and zero crossings is not imposed, as extrema cannot be properly defined for multivariate signals [27]. Filter banks represent an array of band-pass filters designed to isolate input signal into different frequency bands. In this section, we investigate the filter bank property of MEMD for white noise.

Figure 4 shows the frequency response and the corresponding filter bank property of EMD and MEMD for white Gaussian noise. In this experiment, we use 12-channel white Gaussian noise with each of length 5000. The power spectra of its resulting first 9 IMFs (of channel 1) via standard EMD are plotted in Figure 4(a). Next, the same 12-channels data are processed via MEMD and the estimated power spectra of its IMFs (of channel-1) are shown in Figure 4(b). It is observed that the overlapping regions of frequency bands of different IMFs are reduced in the case of MEMD as compared with standard EMD. The alignment of frequency bands for individual IMF in case of MEMD provides much better results than standard [EMD] as illustrated in Figure 4. The results of EMD- and MEMD-based decomposition for financial data of channel 1 with an uncorrelated white noise of same length as that of the input signal are shown in Figure 5 and we observe that larger number of IMFs is obtained by MEMD than the standard EMD. In standard EMD, mode mixing is evident in IMFs where either multiple modes are present or a single mode is leaked into IMFs as compared with MEMD. The original signal, reconstructed signal, and error (of the order $10 \times e^{-13}$) are shown in Figure 5(b). The signal-to-noise ratio (SNR) of the reconstructed signal (with respect to original one) is 318 dB and this SNR proves that the original data is reconstructed simply by summing up the bases with a negligible reconstruction error.

The frequency response is illustrated by MEMD on 5 channel of financial data. The power spectra of its resulting 6 IMFs (of channel-1) are plotted in the top of Figure 6(b). The same financial time series (channel-1) is decomposed using standard EMD and the estimated power spectra of its 6 IMFs are shown in Figure 6(a). It is observed that the alignment of frequency bands for individual IMF of MEMD provides much better results than standard
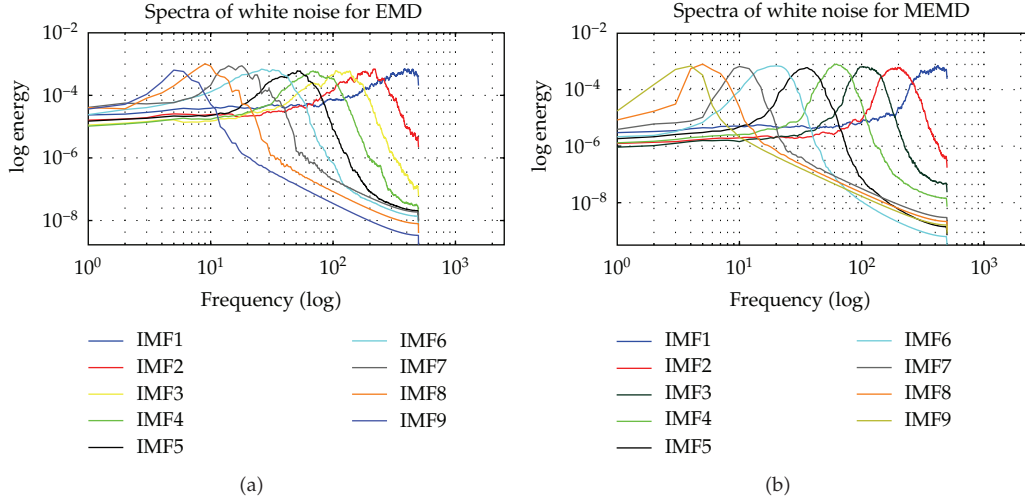
(a)   (b)

Figure 4: Spectra (smoothed) of IMFs (IMF1–IMF9) of channel 1 obtained for 12 channels white Gaussian noise via the standard EMD (a) and MEMD (b). Overlapping of the frequency bands of the IMFs is reduced with MEMD.

EMD. It is already shown in Figure 4 that MEMD of white Gaussian noise properly aligns the bandpass filters associated with the corresponding IMFs than the standard EMD.

3. Prediction Model

In the business and economic situation, it is very significant to predict various kinds of economic variables more precisely in order to expand proper strategies and avoid the risk of potential victims. In the literature, there are some typical techniques that have been developed in predicting financial time series. Among them, most popular technique of linear analysis is perhaps autoregressive moving average (ARMA) method, available in the literature for univariate time series forecasting for stationary time series. The efficiency of the linear method depends on how much the predicted signal is similar to the original signal. The linear prediction model is well fitted with stationary signal and hence prediction error is minimized. When the analyzing signal is decomposed into multiple bands, each band becomes more stationary and holds narrow bandwidth. Then the efficiency of ARMA-based prediction with multiband model is improved. Although wavelet-based decomposition method is treated as a data adaptive technique, it makes the use of priory bases. Whereas EMD-based method is fully data adaptive and the bases are derived from data. The interesting property of EMD technique is that the bases are disjoint in terms of instantaneous frequency at each time instant. With such time-space filtering, IMFs can be filtered without any contamination by the signals from other IMFs. The multivariate EMD is applied here for multiband representation of several channels of financial time series together.

In signal processing, ARMA models are typically applied to autocorrelated time series data. It estimates the future value $s(t)$ with given past values $s(t-1), s(t-2), \ldots\ldots, s(t-k)$. The ARMA (p, q) time series model of signal $s(t)$ is defined as the deterministic relation [28],

$$\tilde{s}(t) = \xi + \sum_{j=0}^{q} \beta_j \varepsilon_{t-j} + \sum_{i=1}^{p} \alpha_i s(t-1), \tag{3.1}$$
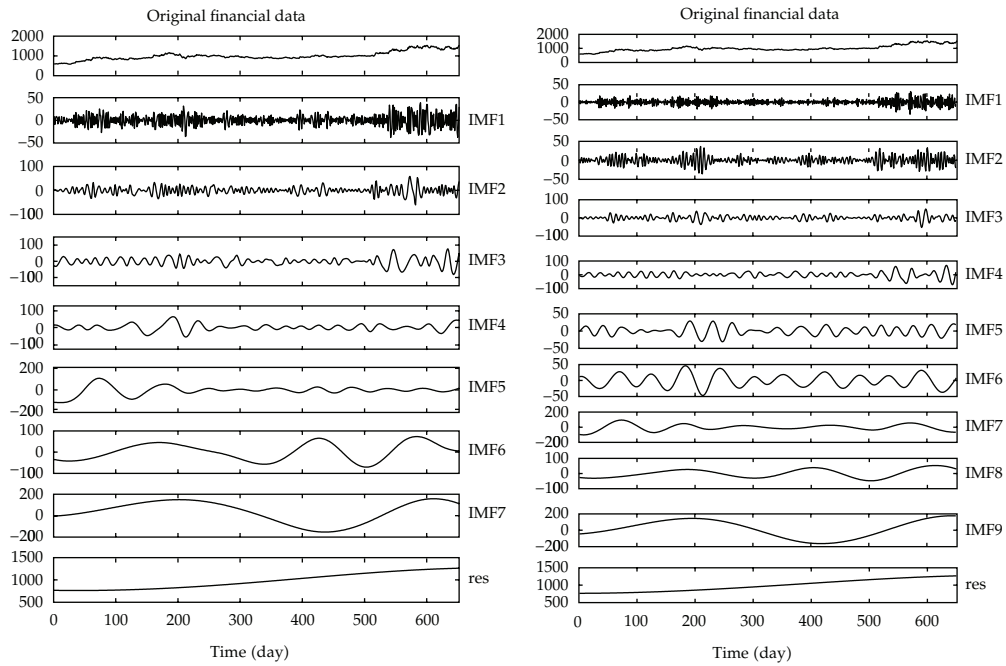
(a)
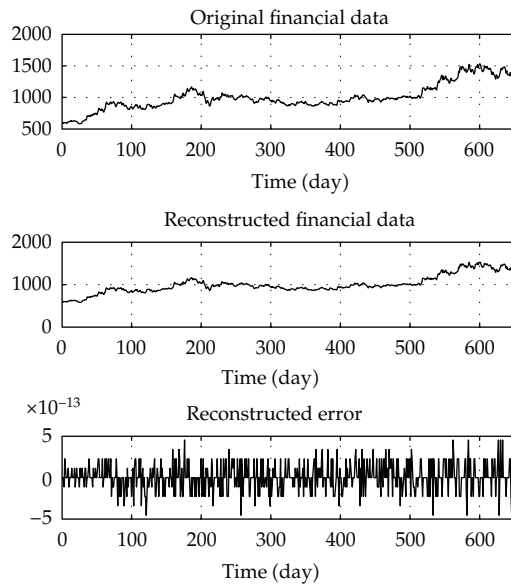
(b)

Figure 5: (a) The results of the EMD and N-A MEMD application for financial data on channel-1. The left and right column represents the IMFs of EMD and MEMD, respectively. (b) Original financial data for channel 1 (top), reconstructed signal by simply adding the IMFs (including the residue) of MEMD (middle) and reconstructed error (bottom). The SNR of the reconstructed signal is 318 dB.
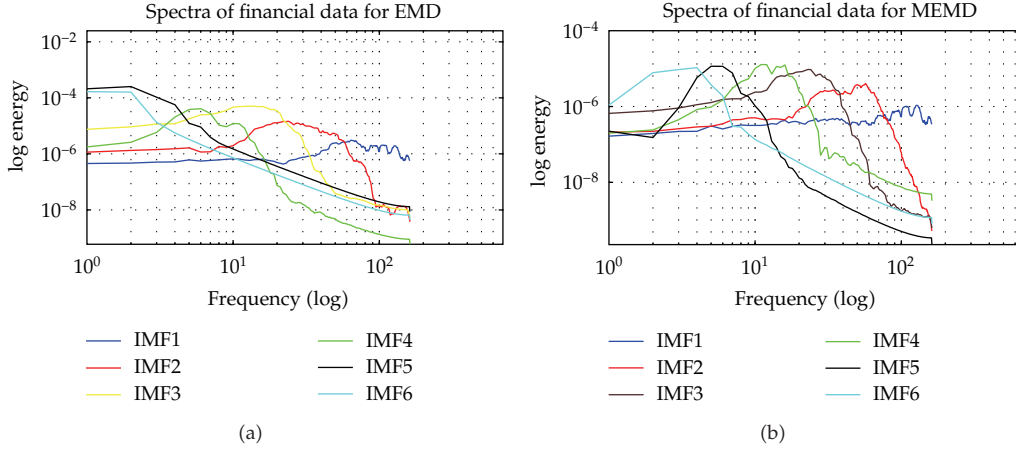
Figure 6: Spectra (smoothed) of IMFs (IMF1–IMF6) of single financial channel obtained by EMD (a) and by MEMD (b). The MEMD is applied on 6-channels (5-channels of financial data and 1 channel of WGN).

where ζ is the intercept, $\sum_{i=1}^{p} \alpha_i s(t-1)$ is autoregressive (AR) part of order p. The order p of AR model tells how many lagged past values are included and $\sum_{j=0}^{q} \beta_j \varepsilon_{t-j}$ is the moving average part (MA) of order q in which the time series is regarded as a moving average of a noise time series ε_t, where $\beta_0, \beta_1, \ldots\ldots, \beta_q$ are parameters of MA model with β_0 fixed as 1 and $\varepsilon_t \sim (0, \delta^2)$ is the white noise with ε_t mutually independent for all t. The objective of prediction is to estimate the future value of the time series as accurately as possible from the current and past values. For a constructed model, the one-step-ahead forecast $\widehat{s}(t)$ given the past can be computed as

$$\widehat{s}(t) = \xi + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} + \sum_{i=1}^{p} \alpha_i s(t-1). \tag{3.2}$$

To implement the multiband prediction model, the predicted signals of individual subband are linearly summed up to reconstruct the overall prediction. Let $\widehat{s}(t)$ be the predicted signal of ith subband $s_i(t)$ denoted by $\widehat{s}_i(t)$ the overall prediction can be obtained as $\widehat{x}(t) = \sum_{i=1}^{M} \widehat{s}_i(t)$ corresponding to the time series $x(t)$, where M is the total number of subbands. It is already mentioned that the original signal is reconstructed with negligible error by summing up the subbands obtained by MEMD. Hence, no reconstruction error will be introduced in the implementation of multiband prediction model. The forecasting error can be determined as the difference between the actual observed value and the one step ahead forecast at any time t as

$$\xi(t) = x(t) - \widehat{x}(t). \tag{3.3}$$

A k-step-ahead AR prediction can be made by recursive application of (3.2). In recursive application, the observed $s(1)$ is used to generate the estimated $\widehat{s}(2)$. That estimate is then substituted as $s(t-1)$ to get the estimated $\widehat{s}(3)$ and so on. Prediction is illustrated in Figure 7 for financial data of channel-2. The observed financial data of length 652 are fitted
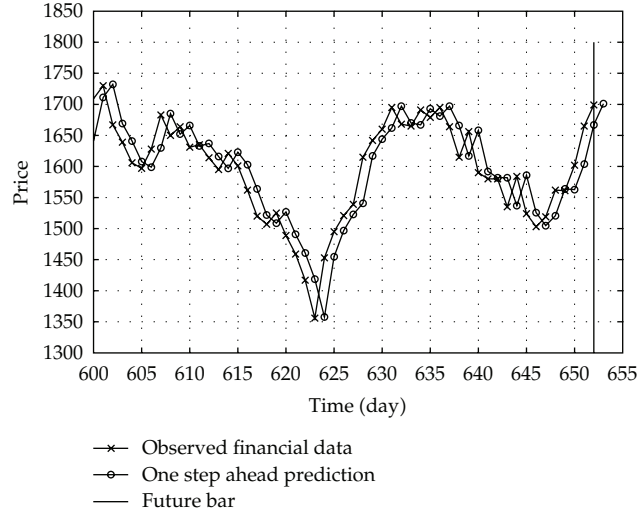
Figure 7: One-step-ahead prediction for single channel (channel-2) financial time series.

with an AR (1) model and the segments of observed and predicted index beginning from 600 is plotted in Figure 7. The observed index ends in 652, and for the days 600–653 the predicted values plotted are one-step-ahead predictions. For the AR (1) model, this means that the prediction for day t is made from observed index in the preceding one day. One-step-ahead predictions in general make use of observed data for times $t \leq (t-1)$ to make the prediction for time t.

4. Performance Analysis

This paper presents several procedures of multichannel financial time series prediction. The important aspect of any prediction procedure is the properties of the error measure that is employed, and whether those properties are appropriate with certain objectives. The difference between the original and reconstructed signals is termed as residual error. In this section we will examine the performance of the proposed multiband ARMA model in terms of SNR, MSE, and rose histogram of the absolute differences of the computed and the actual outputs.

4.1. Signal-to-Noise Ratio (SNR)

The signal-to-noise ratio (SNR) is a good measure of distortion in time domain between original and reconstructed signal. The global SNR values of signal $x(t)$ are determined by the following equation:

$$\text{SNR} = 10 \log_{10} \left[\frac{\sum_{t=0}^{T-1} x^2(t)}{\sum_{t=0}^{T-1} |x(t) - \widehat{x}(t)|^2} \right], \tag{4.1}$$

where $x(t)$ and $\hat{x}(t)$ are the original observed and predicted financial data, respectively, and T is the total number of samples include in data set. This operation is called as global SNR because the summation is performed over the whole data length and expressed as a ratio or factor in units of decibels (dB).

4.2. Mean Square Error (MSE)

In statistics and mathematical world, the mean squared error (MSE) is one of many ways to measure the error rate between the predicted value and the actual data set in time domain. MSE is a risk function, corresponding to the expected value of the squared error loss or quadratic loss. The MSE between actual and predicted value is defined as

$$\text{MSE} = \frac{1}{T} \sum_{t=0}^{T-1} \{x(t) - \hat{x}(t)\}^2, \tag{4.2}$$

where $\hat{x}(t)$ is considered the predicted value and $x(t)$ is the actual data set value to be predicted. The measure of MSE is a direct time domain method for performance evaluation of ARMA algorithm. The objective of any predicted system is to minimize this MSE.

5. Experimental Results

It is well known that most statistical models are restricted on a static character. The quantified relationships and system parameters remain the same across the observation and forecasting time period. Therefore, the ability to apply and test more flexible and sophisticated algorithms in order to achieve better forecasting performance seems to be very promising. ARMA models show a relative good performance in case of time series with a strong autoregressive character. In this experiment, the financial data are obtained from Clayon Financial, Japan representing daily average pricing of different oil companies. The maximum, minimum, opening, closing pricing data are not available though they are important factors. This work is confined to only the prediction of next day pricing value and it is also possible to extend this work for multistep prediction with trend analysis. The financial time series of five different but homogeneous companies (termed here as five channels) as shown in Figure 8 are studied to evaluate the performance of the proposed method.

The time series is decomposed into multiple subbands using discrete wavelet transform (DWT), EMD, and multivariate EMD. The EMD and MEMD are being fully data adaptive, they need not require any input parameter except the stopping criteria what is set to 0.25. The average number of IMFs obtained by using classical EMD is 7, whereas the number is 10 for MEMD. The bi-orthogonal wavelet basis db8 is used to decompose the data channels into four subbands. It is noted (experimentally found) that the more the number of sub-bands with DWT produces higher reconstruction error. The ARMA parameters are computed for each subband and the overall prediction is obtained by summing up the predicted data from each subband. It is noted that the subbands obtained by applying EMD or MEMD are fully data adaptive without employing any predefined basis function. The original data and predicted data are shown in Figure 9 in which it is observed that all the predicted data are almost superimposed to original time series.
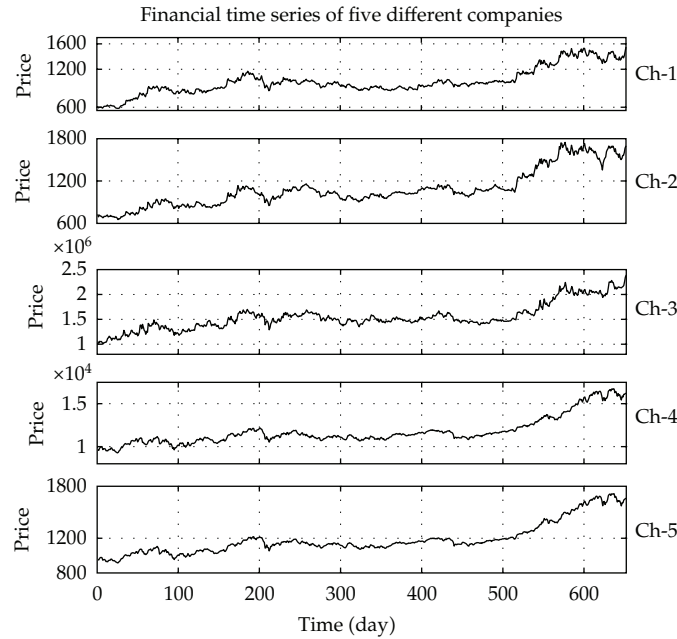
Figure 8: Financial time series of five different (but homogeneous) companies.

The EMD method decomposes a signal based on its intrinsic time scales. The energy of the IMFs with different local scales is distributed in different frequency bands. It is already mentioned in the above for white noise that the original EMD acts as a dyadic filter bank, similar to the dyadic wavelet transform [23].

In this aspect, the research results are presented here on the behavior of the MEMD. We have taken five channels of financial time series. The whole financial data space is decomposed into IMFs using MEMD. Using a k step-ahead AR prediction model, the original IMFs (for channel-2) and the corresponding predicted signals are shown in Figure 10. It is observed that the predicted signals (blue) are very close to the original signal at higher-order IMFs, whereas error is increased with decreasing the order of IMFs, that is, more prediction errors occurred for the IMFs with higher bandwidth. After applying MEMD, the IMFs are transformed to frequency domain using FFT. Then, the power spectrum is estimated for each mode as the average of the squared absolute values of the corresponding Fourier transforms over all the channels. The average spectra are shown in Figure 11 in which it is observed that in locally, each IMF contains lower frequency oscillation than the one extracted just before. This means that each higher-order IMFs contain lower bandwidth than that of lower-order IMFs. That is why the prediction performance is increased for higher order IMFs.

The proposed algorithm is also evaluated in the financial time series with SNR, MSE, and absolute difference between the actual and predicted values. The results are summarized in Figures 12, 13, 14, and 15. In Figures 12 and 13, the rose histogram of the absolute differences between the actual and the estimated values on the test set, as well as the plot of the actual and the predicted values, is given for each of the four different approaches. It is clearly observed from Figures 12 and 13 that the distribution of values is grouped according to their numeric range. Each group is shown as one bin. The length of each bin reflects the number of elements of absolute differences of the computed and the actual outputs in an
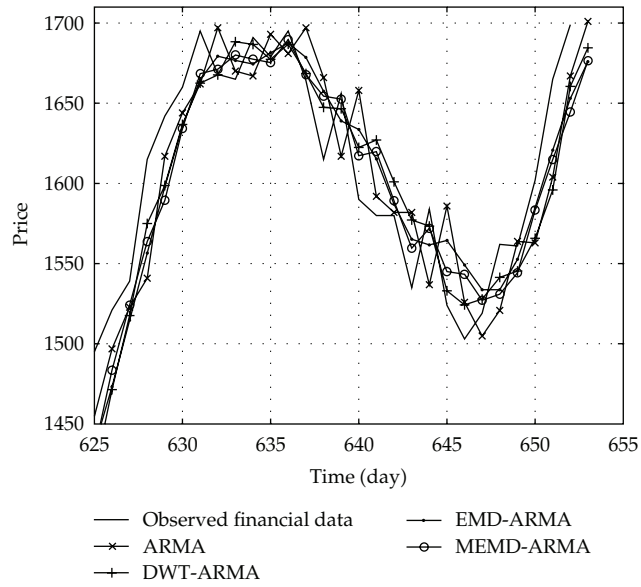
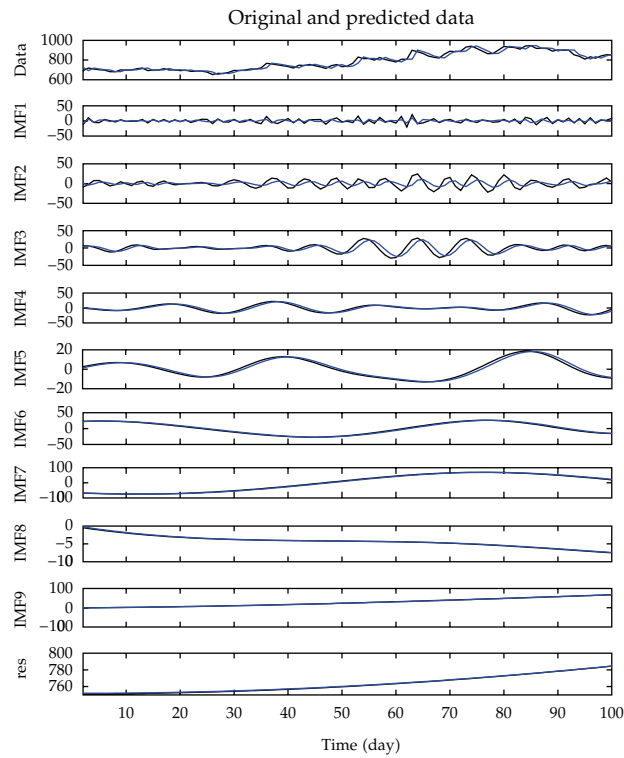**Figure 9:** Original time series and predicted data using different model.

Figure 10: The original financial time series (black) as well as its IMFs (obtained by applying MEMD) for channel-2 and the corresponding predicted signals (blue).
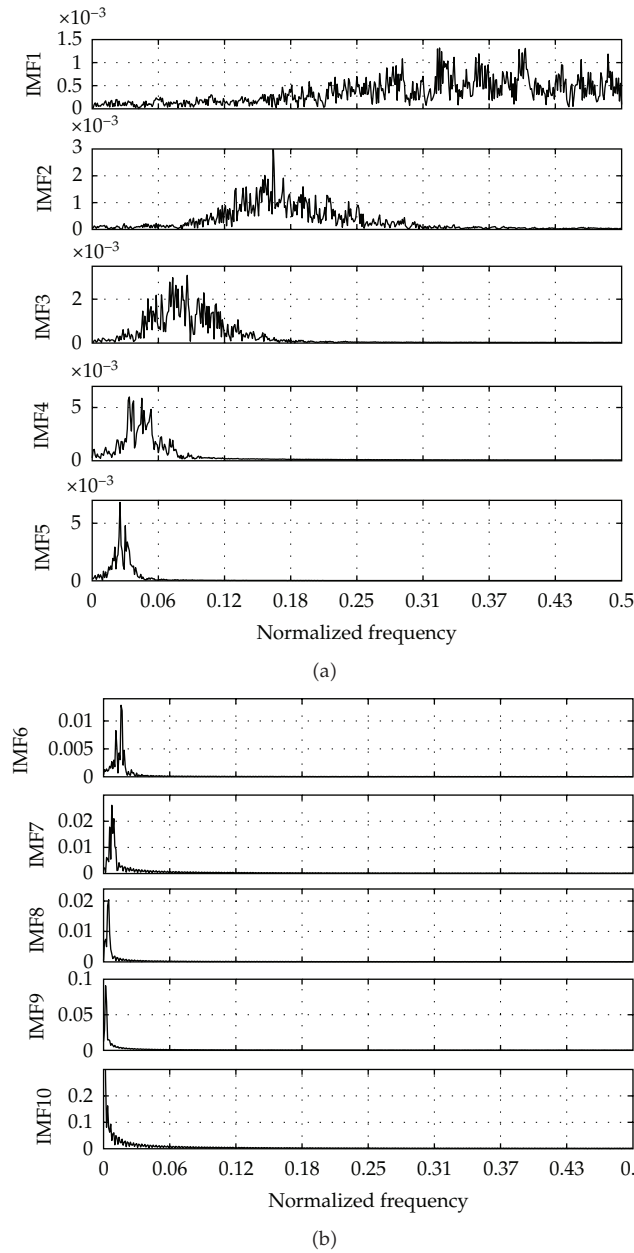
Figure 11: The average spectrum of individual IMFs over five channels of financial time series; (a) IMF1–IMF5, (b) IMF6–IMF10.

angle that falls within a group, which ranges from 0 to the greatest number of elements deposited in any one bin and it is observed graphically that MEMD provides better result in comparison with wavelet and EMD-based multiband predictions.

The performance in terms of SNR in dB and MSE corresponding to the EMD, MEMD, and wavelet-based multiband methods is shown in Figures 14 and 15, respectively. It is observed from Figures 14 and 15 that the SNR with MEMD-based method is always higher
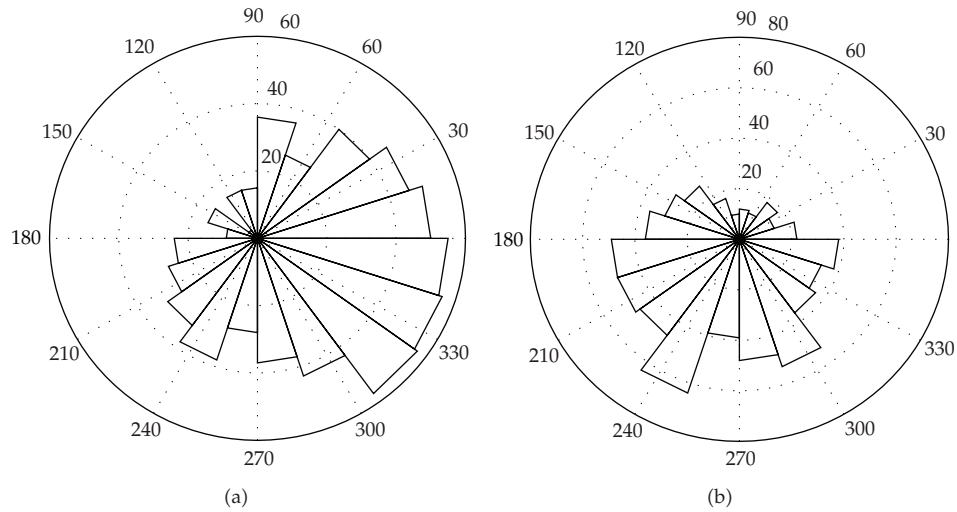
Figure 12: The rose histogram of elements of the absolute differences of the predicted and original data using ARMA (a) and DWT-ARMA (b) for single channel (ch-2).
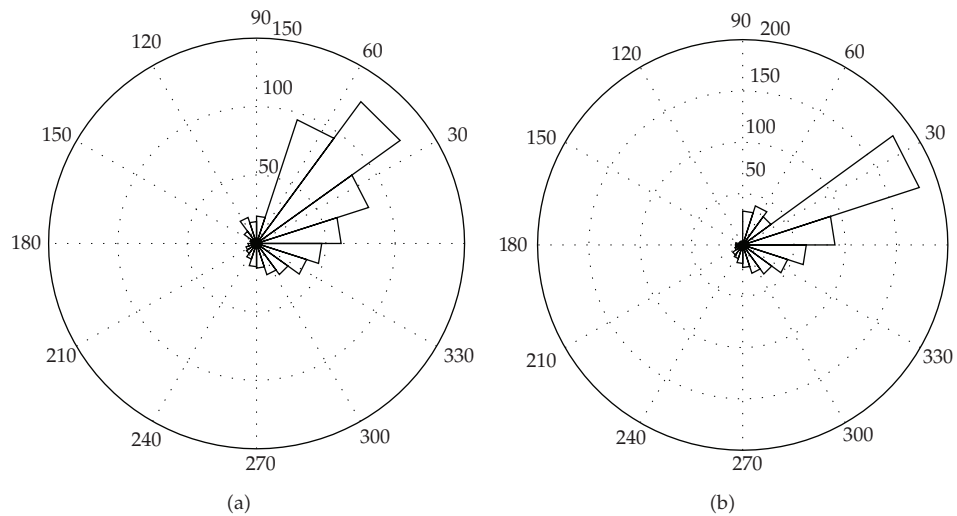

Figure 13: The rose histogram of elements of the absolute differences of the predicted and original data using EMD-ARMA (a) and MEMD-ARMA (b) for single channel (ch-2).

than the EMD and wavelet-based multiband ARMA model, and MSE is always smaller than wavelet and EMD. It is already showed that original data is reconstructed using DWT, EMD, and MEMD with a negligible error. All of these criteria explain that the MEMD gives more effective prediction results and better results than DWT and EMD-based ARMA forecasting model.
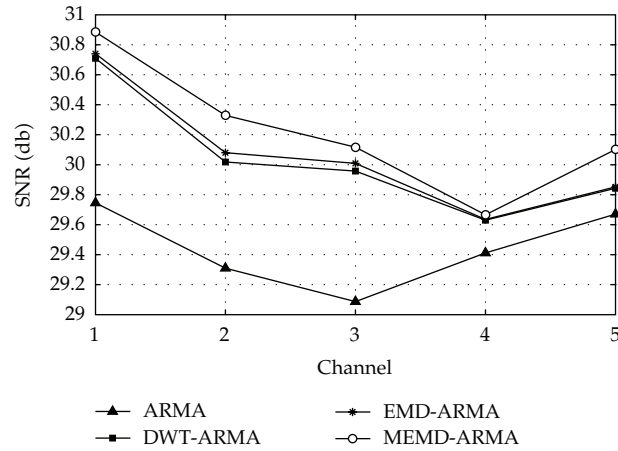
Figure 14: Comparison (in SNR) of proposed MEMD-ARMA, model with ARMA, DWT-ARMA and EMD-ARMA for different channels of financial data.
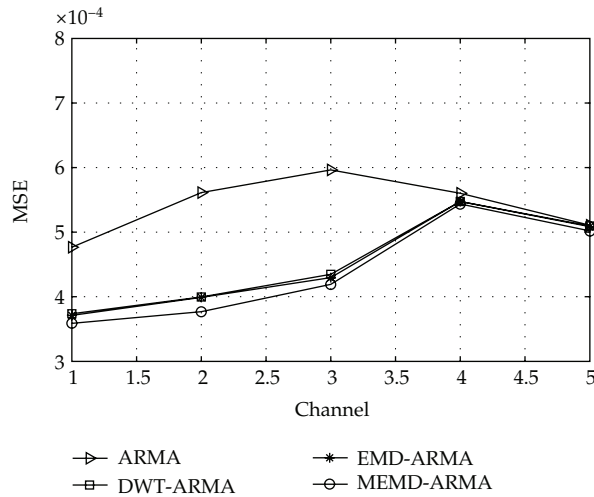

Figure 15: Comparison (in MSE) of proposed MEMD-ARMA model with ARMA, DWT-ARMA, and EMD-ARMA for different channels of financial data.

6. Discussion and Conclusions

A novel method of multiband prediction model of financial data is implemented with multivariate empirical mode decomposition (MEMD) which is a time domain, data adaptive filter bank analysis scheme. The predicted data using the proposed model are closest to the original time series.

The Fourier transform uses priori sinusoidal bases and is suitable only for stationary signals. The wavelet transform has already become popular as a data adaptive multiband decomposition method using priori wavelet basis and hence limited in some extend for non-stationary signals. The EMD-based multiband representation technique is based on the local characteristic time scale of the data. Its bases (IMFs) used to represent the given signal are nonlinear functions that are directly extracted from the data. Although EMD-based

decomposition method is a fully data adaptive method, its mode mixing is evident between intrinsic mode functions (IMFs) or a single mode is "leaked" into two IMFs. In BEMD and trivariate EMD, number of channel selection is crucial issue from multiple channels signal and trivariate EMD takes projections along multiple directions in three-dimensional spaces. Whereas the MEMD process straightforwardly the multichannel input data in n-dimensional spaces without any question of channel selection. It aligns better with the corresponding IMFs from different channels across the same frequency range and effects of mode mixing have been reduced which is crucial for real-world applications which are discussed above. Each IMF is a monocomponent in the local scale and hence properly modeled by ARMA parameters reducing the prediction error. The MEMD works better with the homogeneous data for different channels. In this experiment, the financial time series of the companies similar in nature (different oil companies) are used. Moreover, it produces more IMFs with narrow bandwidth than classical EMD. The increased number of IMFs with narrow bandwidth is well fitted with ARMA model and hence better prediction is achieved using the proposed MEMD-ARMA model.

The proposed method has been compared with wavelet-based subband decomposition method using biorthogonal wavelet. It should be noted that biorthogonal wavelet has the capability to reconstruct the original signal with minimum error compared to the other wavelets. In wavelet-based filter bank, it is a critical decision to select the proper wavelet function for decomposition based on the characteristics of analyzing signals. In the EMD and MEMD-based methods, no input parameter is required for decomposition and perfect reconstruction is obtained by simply adding the IMFs (also the residue), hence the proposed one is superior to the other multiband ARMA models. The residual error of EMD-based multiband ARMA method is always less than the errors introduced by wavelet-based subband ARMA model. The specialty of EMD-based multiband method is that it is full data adaptive decomposition without approximation of signal components based on the priori basis functions.

Generally, the result of the approximation series under the wavelet transforms; EMD and MEMD is better than the original return data and more stable in variance, mean, and no outliers. Furthermore, the forecasting using ARMA under multiband decomposition is better than forecasting directly, and also it gives more accurate results. Finally, all the mentioned methods are compared (in terms of SNR, MSE, and rose histogram of the absolute differences of the predicted and the actual data). It is found that the MEMD-based approach performs better because of its full data adaptive in nature and alignment of the corresponding IMFs obtained from different homogeneous time series across the same frequency range.

References

[1] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis*, Prentice Hall, Englewood Cliffs, NJ, USA, 3rd edition, 1994.

[2] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, NJ, USA, 1996.

[3] G. J. Deboeck, *Trading on the Edge: Neural, Genetic and Fuzzy Systems for Chaotic Financial Markets*, John Wiley & Sons, New York, NY, USA, 1994.

[4] Y. S. Abu-Mostafa and A. F. Atiya, "Introduction to financial forecasting," *Applied Intelligence*, vol. 6, no. 3, pp. 205–213, 1996.

[5] S. Rao and W. A. Pearlman, "Analysis of linear prediction, coding, and spectral estimation from subbands," *IEEE Transactions on Information Theory*, vol. 42, no. 4, pp. 1160–1178, 1996.

[6] A. Härmä, "Linear predictive coding with modified filter structures," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 8, pp. 769–777, 2001.

[7] R. A. Finan, R. I. Damper, and A. T. Sapeluk, "Improved data modeling for text-dependent speaker recognition using sub-band processing," *International Journal of Speech Technology*, vol. 4, no. 1, pp. 45–62, 2001.

[8] E. Erzin, A. Enis Cetin, and Y. Yardimci, "Subband analysis for robust speech recognition in the presence of car noise," in *Proceedings of the 20th International Conference on Acoustics, Speech, and Signal Processing (ICASSP '95)*, vol. 1, pp. 417–420, May 1995.

[9] W. C. Chen, C. Hsieh, and E. Lai, "Multiband approach to robust text-independent speaker identification," *Journal of Computational Linguistics and Chinese Language Processing*, vol. 9, no. 2, pp. 63–76, 2004.

[10] S. L. Tan and T. R. Fischer, "Linear prediction of subband signals," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 9, pp. 1576–1583, 1994.

[11] N. E. Huang and S. S. P. Shen, *Hilbert-Huang Transform and Its Applications*, vol. 5 of *Interdisciplinary Mathematical Sciences*, World Scientific, Hackensack, NJ, USA, 2005.

[12] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society A*, vol. 454, no. 1971, pp. 903–995, 1998.

[13] A. H. Tewfik, "Potentials and limitations of wavelets in signal acquisition and processing," in *Proceedings of the 2nd International IEEE EMBS Conference*, pp. 328–329, 1993.

[14] S. Alonso-Quesada and M. De la Sen, "Robust adaptive stabilization of linear time-invariant dynamic systems by using fractional-order holds and multirate sampling controls," *Discrete Dynamics in Nature and Society*, vol. 2010, Article ID 620546, 27 pages, 2010.

[15] M. De la Sen, "On Chebyshev's systems and non-uniform sampling related to Caputo fractional dynamic time-invariant systems," *Discrete Dynamics in Nature and Society*, vol. 2010, Article ID 846590, 24 pages, 2010.

[16] P. Flandrin, G. Rilling, and P. Gonçalvés, "Empirical mode decomposition as a filter bank," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 112–114, 2004.

[17] T. Tanaka and D. P. Mandic, "Complex empirical mode decomposition," *IEEE Signal Processing Letters*, vol. 14, no. 2, pp. 101–104, 2007.

[18] N. Rehman and D. P. Mandic, "Empirical mode decomposition for trivariate signals," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1059–1068, 2010.

[19] L. Hong, "Decomposition and forecast for financial time series with high-frequency based on empirical mode decomposition," *Energy Procedia*, vol. 5, pp. 1333–1340, 2011.

[20] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, Calif, USA, 1998.

[21] A. Cohen, *Numerical Analysis of Wavelet Methods*, vol. 32 of *Studies in Mathematics and Its Applications*, North-Holland, Amsterdam, The Netherland, 2003.

[22] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proceedings of The Royal Society of London Series A*, vol. 466, no. 2117, pp. 1291–1302, 2010.

[23] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proceedings of the Royal Society A*, vol. 460, no. 2046, pp. 1597–1611, 2004.

[24] G. Rilling, P. Flandrin, P. Goncalves, and J. M. Lilly, "Bivariate empirical mode decomposition," *IEEE Signal Processing Letters*, vol. 14, no. 12, pp. 936–939, 2007.

[25] N. Rehman and D. P. Mandic, "Filter bank property of multivariate empirical mode decomposition," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2421–2426, 2011.

[26] N. E. Huang, Man-L.C. Wu, S. R. Long et al., "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis," *The Royal Society of London. Proceedings Series A*, vol. 459, no. 2037, pp. 2317–2345, 2003.

[27] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*, John Wiley & Sons, New York, NY, USA, 2009.

[28] B. Thiesson et al., "ARMA time series modeling with graphical model," in *Proceedings of International Conference on Uncertainty in Artificial Intelligence*, pp. 552–560, 2004.