The Korean Journal of Applied Statistics (2010)
23(2), 249–261

Long-Term Forecasting by Wavelet-Based Filter Bank Selections and Its Application

Jeong-Ran Lee[1] · Youlim Lee[2] · Hee-Seok Oh[3]

[1]Department of Statistics, Seoul National University
[2]Cooperative Banking Support Department, NH Bank
[3]Department of Statistics, Seoul National University

(Received January 2010; accepted March 2010)

Abstract

Long-term forecasting of seasonal time series is critical in many applications such as planning business strategies and resolving possible problems of a business company. Unlike the traditional approach that depends solely on dynamic models, Li and Hinich (2002) introduced a combination of stochastic dynamic modeling with filter bank approach for forecasting seasonal patterns using highly coherent(High-C) waveforms. We modify the filter selection and forecasting procedure on wavelet domain to be more feasible and compare the resulting predictor with one that obtained from the wavelet variance estimation method. An improvement over other seasonal pattern extraction and forecasting methods based on such as wavelet scalogram, Holt-Winters, and seasonal autoregressive integrated moving average(SARIMA) is shown in terms of the prediction error. The performance of the proposed method is illustrated by a simulation study and an application to the real stock price data.

Keywords: Filter bank, High-C waveforms, long-term forecasting, scalogram, time series analysis, wavelets.

1. Introduction

When analyzing a time series a classic problem has been its decomposition into components: the trend, the seasonal component, and the irregular term, since, frequently, it can be easier to forecast these components of a time series than the whole series itself. In economic time series, the seasonal component has usually a constant period of 12 months and to assess it one uses some underlying assumptions or theory about the nature of the series. Long-term trend means the fluctuation of a time series on time scales of more than one year. It can be considered the long-term trend of the series changes with the pulse of the general economy. Such a component is found by elimination of the seasonal component and the irregular term. Since this component is usually related to general economic fluctuations, it is often referred to as a "business cycle". In the meanwhile, seasonality means the inherent periodicity in many economical time series. Periodic or seasonal time series

This work was supported by the Korea Research Foundation Grant funded by the Korean Government(MOEHRD)(KRF-2007-313-C00107).

[3]Corresponding author: Associate Professor, Department of Statistics, Seoul National University, 599 Gwanak-ro, Gwanak-gu, Seoul 151-742, Korea. E-mail: heeseok@stats.snu.ac.kr

are not deterministic because there are considerable random variations in the waveforms and so they do not perfectly repeat themselves from one period to another. Forecasting such changing seasonal time series is therefore difficult and there are many possible ways to model the periodicity or seasonality depending on how the time series evolve over time.

We focus on long-term forecasting of seasonal time series via wavelet methodology. Research on long-range forecasting methods which are applicable to all areas of the social, behavioral, and management sciences has been growing rapidly. Because of the diversity of applications, the literature on this topic is broadly scattered in a large number of journals, including those in fields such as astronomy, economics, engineering, physics, and statistics. However, wavelet methods have not been used extensively in forecasting economic time series though reviews for the role of which in the analysis of statistical time series are noticeable, see, *e.g.*, McCoy and Walden (1996), Nason and von Sachs (1999) and Nason and Silverman (1995). The goal of this paper is to implement the wavelet methodology in detecting cycles and seasonal patterns in time series and then impose a parametric modeling on those extracted patterns to derive the forecasts of the next periods.

There are many possible ways to model changing seasonality such as the seasonal autoregressive integrated moving average(SARIMA) model or multivariate autoregressive moving average(ARMA) model. Since the SARIMA model has the lack of power for reliable long-term forecasting and a full multivariate ARMA model involves modeling complexity, a novel technique to overcome these drawbacks is proposed by Li and Hinich (2002), called a filter bank approach. They introduced a different methodology for modeling and forecasting seasonal time series by combining the pattern-extraction and dimension-reduction technique with parametric dynamic models. The method utilized the highly coherent(High-C) seasonal components to capture long-term seasonal effects and vector autoregressive(VAR) model to obtain the predictor of the components. However, there were some technical ambiguities to implement their method in multiresolution (multiscale) analysis. So we modify the filter selection and forecasting procedure on wavelet domain to be more feasible by using the standard discrete wavelet transform(DWT). We adopt the splitting idea from the wavelet scalogram approach based on wavelet variance estimation methods proposed by Ariño *et al.* (2004). An application of the original scalogram method to the short-term forecasting of corporate bond yields was introduced by Lee (1998), and it was shown to be very effective using wavelet based methods on an analysis of economic time series. In this paper, the performance of the proposed method is illustrated by a simulation study and an application to the real stock price data with a comparison of the results from the existing methods such as Holt-Winters and SARIMA models.

This paper is organized as follows: the general formulation and some theoretical properties of the filter bank approaches with a brief review of wavelets are given in Section 2. In Section 3, our modified method for the filter selection procedure is described in detail. Forecasting with VAR model for the selected waveforms is illustrated in Section 4. A simulation study and a real example result is given in Section 5. Finally, we present some discussions in Section 6 followed by concluding remarks in Section 7.

2. Filter Bank Approach via Wavelets

In this section, we review the decomposition of a seasonal time series by an statistical time-frequency analysis(STFA) filter bank following the approach proposed by Li and Hinich (2002). When the filters are applied to each seasonal vector of a seasonal time series, the resulting projections can be regarded as waveforms that propagate from one period to another. Among these waveforms, Li and Hinich (2002) proposed to model only with the High-C waveforms over period. We briefly introduce the basic formulation for detecting High-C waveforms within a framework of STFA and give an overview of wavelet analysis for the use of wavelet filter banks.

2.1. Decomposition by filter banks

Following the notion of filter bank decomposition in Li and Hinich (2002), one can capture the desirable features of the input time series. Here are some definitions and assumptions used in the context of filter bank approach.

Let $\{y_t,\ t = 0, \pm 1, \pm 2, \ldots\}$ be the (univariate) time series of interest. Suppose that $\{y_t\}$ has observed over n periods of length $p > 1$ so that the available data can be expressed as

$$\boldsymbol{y}(m) \equiv (y_{mp-1}, \ldots, y_{mp-p})^T, \qquad m = 1, \ldots, n,$$

where $\boldsymbol{y}(m)$ is the seasonal vector of the m^{th} period. The first element of $\boldsymbol{y}(m)$ is the most recent observation in period m. Let $\boldsymbol{w}_k \equiv [w_k(1), \ldots, w_k(p)]^T$, $k = 1, \ldots, p$, be a set of orthonormal finite-impulse-response filter, satisfying

$$\boldsymbol{w}_j^T \boldsymbol{w}_k = \begin{cases} 1, & \text{if } j = k, \\ 0, & \text{if } j \neq k. \end{cases}$$

The output from the k^{th} filter $\boldsymbol{w}_k$ with $\boldsymbol{y}(m)$ as an input can be expressed as

$$\xi_k(mp) = \sum_{\tau=1}^{p} w_k(\tau) y_{mp-\tau} = \boldsymbol{w}_k^T \boldsymbol{y}(m).$$

If the samples are collected across the filter bank to form a vector $\boldsymbol{\xi}(m) \equiv [\xi_1(mp), \ldots, \xi_p(mp)]^T$, then it follows that

$$\boldsymbol{\xi}(m) = \boldsymbol{W} \boldsymbol{y}(m),$$

where $\boldsymbol{W}$ is an orthonormal matrix defined by

$$\boldsymbol{W} \equiv [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p]^T.$$

Because $\boldsymbol{W}^T \boldsymbol{W} = \boldsymbol{I}$, one can reconstruct $\boldsymbol{y}(m)$ from $\boldsymbol{\xi}(m)$ according to $\boldsymbol{y}(m) = \boldsymbol{W}^T \boldsymbol{\xi}(m)$. For each fixed k, the seasonal vector $\boldsymbol{y}(m)$ is a superposition of p waveforms $\boldsymbol{w}_k$, whose coefficients $\xi_k(mp)$ vary randomly from one period to another, thus resulting in changing seasonality.

In this decomposition, the intraperiod characteristics of $\{y_t\}$ are encoded by the filters into the projection coefficients, and the interperiod variations of $\{y_t\}$ are transformed into the dynamics of the projection coefficients.

2.2. Detection of High-C waveforms

Once coefficients of filters are obtained, the next step is to detecting High-C waveforms that are most suitable for long-term forecasting. Let us assume that for each k, $\{\xi_k(mp)\}$ is a second-order stationary process with mean μ_k and variance σ_k^2 such that $n^{-1} \sum_{m=1}^{n} \xi_k(mp) \to \mu_k$ and $n^{-1} \sum_{m=1}^{n} \{\xi_k(mp) - \mu_k\}^2 \to \sigma_k^2$ as $n \to \infty$ in probability or almost surely. This is equivalent to assuming that $\{\boldsymbol{y}(m)\}$ is second-order stationary and ergodic. Assuming $\mu_k^2+\sigma_k^2>0$, the coherence of the k^{th} waveform of $\boldsymbol{y}(m)$ is defined as

$$\gamma_k \equiv \frac{\mu_k^2}{\mu_k^2+\sigma_k^2}.$$

The k^{th} waveform of $\boldsymbol{y}(m)$ is said to be (a) completely coherent if $\gamma_k = 1$, (b) incoherent if $\gamma_k = 0$, (c) partially coherent if $0 < \gamma_k < 1$. Let $\hat{\gamma}_k$ be an estimate of γ_k from a finite sample. According to Proposition 1 given in Li and Hinich (2002), $\boldsymbol{w}_k$ is said to be highly coherent, or High-C, if $\hat{\gamma}_k > \theta$, where $\theta \equiv \chi^2_{1-\alpha}(1)/(n+\chi^2_{1-\alpha}(1))$, $\chi^2_z(1)$ is the cumulative distribution function of $\chi^2(1)$, and $\alpha \in (0,1)$ is the significance level of the hypothesis testing problem whether $\mu_k = 0$. In their paper, Li and Hinich (2002) utilize only the High-C waveforms in the parametric modeling step.

2.3. Wavelets

In principle, any orthonormal transformation of a seasonal time series is a valid filter bank. However, the use of wavelet filter banks is motivated by their success in applications such as data representation and data compression. Wavelet filters allow the time scale to vary across the filters and thus gives a grand view of a seasonal vector as well as smaller fractions for a closer look. Li and Hinich (2002) emphasize the unique interpretations of wavelet filter banks to extract the High-C waveforms. However, we review the basics of wavelet transform using a more general filter notation given in Nason and Silverman (1995) that is useful subsequently in this paper to modify the filter selection method of Li and Hinich (2002). For mathematical and practical aspects of wavelets, see, *e.g.* Daubechies (1992) and Vidakovic (1999).

Wavelets are based on functions called scaling functions ϕ which have two key properties. Firstly, $\phi(t)$ and all its integer translates $\phi(t+j)$ form an orthonormal set in L^2, so that $\int \phi(t)^2 dt = 1$ and $\int \phi(t)\phi(t+j)dt = 0$ for integers $j \neq 0$. Secondary, ϕ can be expressed as a linear combination of half-integer translates of itself at double the scale:

$$\phi(t) = \sqrt{2}\sum_j h_j \phi(2t-j).$$

Defining the sequence $\{g_j\}$ as the mirror of the sequence $\{h_j\}$ by the relation of $g_j = (-1)^j h_{1-j}$, the mother wavelet ψ is then defined by

$$\psi(t) = \sqrt{2}\sum_j g_j \phi(2t-j),$$

which is called the dilation equation. Here the sequence notation $\{h_j\}$ or $\mathcal{H}$, a low pass filter and $\{g_j\}$ or $\mathcal{G}$, a high pass filter that both obey the internal and mutual orthogonality relations are called quadrature mirror filters. Bases of various function spaces can now be constructed from appropriate dilations and translations of ϕ and ψ, namely

$$\psi_{j,k}(t) = 2^{\frac{j}{2}}\psi(2^j t - k), \qquad \phi_{j,k}(t) = 2^{\frac{j}{2}}\phi(2^j t-k), \quad j,k \in \mathbb{Z}.$$

Let $\boldsymbol{y} = (y_0, y_1, \ldots, y_{N-1})^T$ be a data vector of length $N = 2^J$, $J \in \mathbb{Z}$. For each $j = 0, 1, \ldots, J-1$ and $k = 0, 1, \ldots, 2^j - 1$, define the DWT with respect to ψ as

$$d_{j,k} = \sum_{t=0}^{N-1} y_t \psi_{j,k}\left(\frac{t}{N}\right),$$

then $d_{j,k}$ is called a wavelet coefficient. If we put the $d_{j,k}$'s be elements of a vector $\boldsymbol{d}$ with respect to indices j and k, then we can write $\boldsymbol{d} = \boldsymbol{W}\boldsymbol{y}$, where $\boldsymbol{W}$ is an orthonormal $N \times N$ matrix. In practice, one can find the DWT coefficients by fast filtering algorithms based on the quadrature mirror filters without actually exhibiting the matrix $\boldsymbol{W}$ as follows. Define the smooth at level J, written by c_J to be the original data,

$$c_J = y_t, \quad \text{for} \ \ t = 0, 1, \dots, 2^J - 1.$$

Now, for $j = J-1, J-2, \dots, 0$, recursively define the smooth c_j at level j and the detail d_j at level j by

$$c_j = \mathcal{D}_0 \mathcal{H} c^{j+1} \quad \text{and} \quad d_j = \mathcal{D}_0 \mathcal{G} c^{j+1},$$

where $\mathcal{D}_0$ is the binary decimation operator, which simply chooses every even member of a sequence.

3. Some Modifications

There were some technical ambiguities to implement the original filter bank approach for detecting High-C waveforms in terms of DWT. If we see the building blocks of DWT in Section 2.3, it is noticeable that they only involve filters with dyadic length of coefficients for each level of decomposition. Since we cannot guarantee that p, the length of a period defined in Section 2.1, is a dyadic integer and there are different numbers of coefficients in each level, it is not feasible to present the coherence of waveforms in terms of DWT. So we modify the filter selection procedure on wavelet domain by combining the idea used in wavelet variance estimation method based on wavelet scalogram proposed by Ariño *et al.* (2004). To extract meaningful patterns from a seasonal time series for reliable forecasting, we first split the scales of wavelet coefficients obtained from DWT by long-term trend and short-term seasonality groups (a similar way used in the scalogram approach) based on the coherence plot as follows.

3.1. Wavelet scalogram approach

The wavelet variance estimation method decomposes the variance of a time series into components associated with different scales. Just as the periodogram produces an ANOVA decomposition of the energy of a signal into different Fourier frequencies, this method gives an alternative way of looking at the energy of a time series in the frequency domain.

Since the periodogram is an estimator of the true spectrum of a process, if the periodogram peaks at a certain level then it could be concluded that the component with that frequency accounts for a large part of the variance of the series in its Fourier decomposition. This gives an wavelet analogue interpretation as follows. We can select the filters at certain levels placed above where there exist energy peaks in the wavelet variance decompositions. Then the selected filters are considered to be those that account for the changing seasonality.

Let $\boldsymbol{d} = (c_{00}, d_{00}, d_{10}, d_{11}, d_{20}, \dots, d_{J-1,2^{J-1}-1})^T$ be the vector of the DWT coefficients of $\boldsymbol{y}$ defined as in Section 2.3. The energy component $E(j)$ of the coefficient vector $\boldsymbol{d}$ at each level j is defined as

$$E(j) = \sum_{k=0}^{2^j-1} d_{j,k}^2,$$

for $j = 0, 1, \ldots, J-1$. The scalogram of the data is a vector of energies

$$\left(c_{00}^2, E(0), E(1), \ldots, E(J-1)\right)^T.$$

Ariño *et al.* (2004) proposed methods of splitting the high- and low-frequency components of the DWT coefficients around where two or more peaks in the above scalogram. For example, if one peak is presented in the scalogram at level s, we can split the wavelet coefficients in $\boldsymbol{d}$ into two new wavelet decompositions, $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ in the following way,

$$d_{j,k}^{(1)} = \begin{cases} d_{j,k}, & \text{for } j = 0, 1, \ldots, s-1, \\ 0, & \text{for } j = s, \ldots, J-1 \end{cases}$$

and

$$d_{j,k}^{(2)} = \begin{cases} 0, & \text{for } j = 0, 1, \ldots, s-1, \\ d_{j,k}, & \text{for } j = s, \ldots, J-1. \end{cases}$$

Here, $\boldsymbol{d}^{(1)}$ is used for the reconstruction of long-term trend and $\boldsymbol{d}^{(2)}$ for short-term periodicity. Since Ariño *et al.* (2004) considered the decomposed components only for the use of the estimation of a time series, they just summed the two parts after a reconstruction. However, our objective is extracting seasonal patterns in a given time series and forecasting its evolution over time, we consider the reconstructions of $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ for later use in parametric forecasting models. In order to remove the unwanted noise from the reconstruction with these selected levels components, we pad the DWT coefficients at the largest $(J-1)^{th}$ level in $\boldsymbol{d}^{(2)}$ by zeros. In wavelet regression, we generally use a special type of shrinkage as a denoising technique called thresholding. Various thresholding methods, such as VisuShrink, SureShrink (Donoho and Johnstone, 1994) or EbayesThresh (Silverman, 2005) are well developed and easily applicable. However, these thresholding methods work better than linear shrinkage when finding jumps in a function that is smooth except that it has some jumps in a few places. If we expand this function in a wavelet basis, for example, the coefficients will be sparse, meaning that most will be small except for a few coefficients corresponding to the location of the jumps. For long-term seasonal time series, however, forecasting seasonality and trend based on each of their mutual coherence is more important than just estimating variations at certain locations. Therefore, we choose to kill only the finest level coefficients as a method of denoising and conserve inner structure of all level coefficients without information loss.

3.2. Coherence plot approach

As an analogue of wavelet scalogram approach, we modify the selection of High-C waveforms in Section 2.2 to the split of waveforms by their coherences. High-C waveforms with γ_k's above the threshold θ are grouped together and the others are classified as another group. If the majority of a group is low-frequency level components, it is reconstructed as a long-term trend. And the short-term seasonality is constructed by the components of the other group with replacement of the largest $(J-1)^{th}$ level by zeros. To be accurate, let $J = 9$, for example, and the High-C waveforms are in a group of levels $\{1, 2, 3, 5\}$ then the rest is grouped by $\{4, 6, 7\}$ since the noise level 8 is thresholded. The majority (3 out of 4) of High-C group is low-frequency level components (less than $J/2 = 4.5$), thus the High-C group $\{1, 2, 3, 5\}$ is reconstructed as a long-term trend, and $\{4, 6, 7\}$ is reconstructed as a short-term seasonality. The reason why we can separate the subsets of wavelet coefficients by their grouped levels is that the waveforms are the elements of nested function spaces on a multiresolution framework. Because our focus is on the seasonal patterns, it needs to be assumed that nonseasonal trends have been removed by some detrending techniques (such as differencing or regression). Instead, by using the reconstructed long-term trend part in forecasting together with the extracted seasonality, we simplify the procedure to contain those detrending steps naturally.

4. Forecasting

Now we have only split filters from the methods based on the High-C waveforms and scalogram, respectively. Since we decompose a seasonal time series into components of the trend and the seasonal patterns, and removed the irregular term, it would be easier to model and forecast components of a time series than the whole series itself. We employ the VAR, which is relatively versatile and computationally simple, to model those waveforms' interperiod dynamics following the literature of Li and Hinich (2002). The reconstructed components of the trend and seasonal patterns are modeled as a bivariate VAR(1) to get the predictor of next periods. We will further compare the result with other prediction methods such as Holt-Winters and SARIMA models in the subsequent sections.

The VAR model (Lütkepohl, 2007) is proven to be especially useful for describing the dynamic behavior of economic and financial time series and for forecasting, see, *e.g.*, Brockwell and Davis (1991) and Hamilton (1994). In this section, we review the definitions and assumptions of VAR model in brief.

The VAR(p) process for K variables, $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{Kt})^T$ is defined as

$$\boldsymbol{y}_t = \Phi_1 \boldsymbol{y}_{t-1} + \cdots + \Phi_p \boldsymbol{y}_{t-p} + \boldsymbol{\varepsilon}_t,$$

where Φ_i for $i = 1, \ldots, p$ are $K \times K$ real valued coefficients matrices, and $\boldsymbol{\varepsilon}_t$ is a K-dimensional white noise(WN) process with positive definite covariance matrix, *i.e.*, $\{\boldsymbol{\varepsilon}_t\} \sim \mathrm{WN}(0, \Sigma_p)$. The coefficients of VAR(p) process can be estimated efficiently by a simple estimation procedure based on the multivariate Durbin-Levinson algorithm for fitting autoregressions of increasing order. It is also analogous to the procedure of the univariate case.

Suppose that we observe $\{\boldsymbol{y}_t,\ t = 1, \ldots, T\}$ of a zero mean stationary multivariate time series with some corresponding sample covariance functions. Then the fitted VAR(p) process ($p < T$) is

$$\boldsymbol{y}_t = \hat{\Phi}_1 \boldsymbol{y}_{t-1} + \cdots + \hat{\Phi}_p \boldsymbol{y}_{t-p},$$

where the coefficients $\hat{\Phi}_1, \ldots, \hat{\Phi}_p$ are computed recursively from the multivariate Durbin-Levinson algorithm. The order p of the autoregression may be chosen to minimize some standard criteria, such as Akaike information criteria(AIC) (Brockwell and Davis, 1991),

$$\mathrm{AIC} = \log\left(\det\left(\widehat{\Sigma}_p\right)\right) + \frac{2}{T} n K^2,$$

or HQ (Hannan and Quinn, 1979),

$$\mathrm{HQ} = \log\left(\det\left(\widehat{\Sigma}_p\right)\right) + \frac{2\log(\log(T))}{T} n K^2,$$

where $\widehat{\Sigma}_p = T^{-1} \sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t^T$ with $\hat{\varepsilon}$ is the sample residual for observation t from an ordinary least squares(OLS) regression of y_{it} on a vector containing a constant term and p lags of each of the elements of $\boldsymbol{y}_t$, and n is the lag-order. For more details, readers are directed to Hamilton (1994).

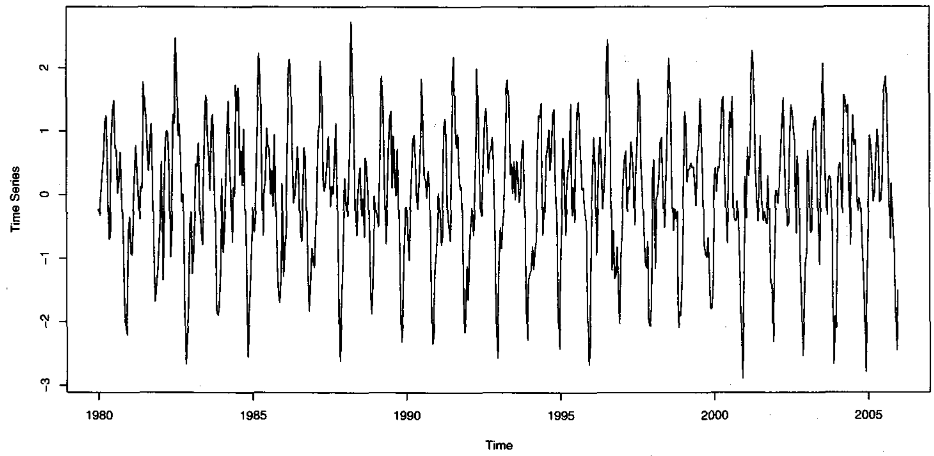

Figure 5.1. A simulated seasonal time series with periodicity $p = 24$.

5. Numerical Experiments and an Application

In this section, we compare the results of different filter bank approaches and forecasting procedure by a simulation study and a real data example. To measure the performance of forecasts up to a given horizon h, we employ the prediction error(PE) defined by

$$\mathrm{PE} = \frac{\sum_{i=1}^{h}\{\hat{y}_t(i) - y_{t+i}\}^2}{h},$$

where $\hat{y}_t(i)$ is the i-step ahead prediction based on $y_1, \ldots, y_t$. The smaller PE indicates the better performance of a given method.

5.1. Simulation study

Before testing the proposed methods on real data, it is necessary to understand the methods under controlled conditions. To this end, a seasonal time series is simulated with period $p = 24$ and three sinusoidal components plus a Gaussian white noise(GWN) component. The amplitudes of the sinusoids remain constant but the phases change randomly from one period to another as independent AR(1) processes. More precisely,

$$x_t = s_t + \varepsilon_t,$$

where $\varepsilon_t \sim \mathrm{GWN}(0, \sigma^2)$,

$$\begin{aligned} s_t &\equiv \sum_{i=1}^{3} \beta_i \cos\{2\pi f_i t + \phi_i(m)\}, \quad t = mp - p, \ldots, mp - 1, \\ \phi_i(m) &\equiv \pi\{b_i \varphi_i(m) - c_i\}, \\ \varphi_i(m) &\equiv a_i \varphi_i(m-1) + e_i(m) \end{aligned}$$

and $e_i(m) \sim \mathrm{GWN}(0, 1 - a_i^2)$. In this experiment, $(\beta_1, \beta_2, \beta_3) = (1, 0.7, 0.7)$, $(f_1, f_2, f_3) = (1/p, 2/p, 4/p)$, $(a_1, a_2, a_3) = (0.8, 0.8, 0.8)$, $(b_1, b_2, b_3) = (0.2, 0.3, 0.4)$ and $(c_1, c_2, c_3) = (1, 0.5, 0.1)$. The value of σ^2 is chosen to 10dB. A realization of x_t, $t = 1, \ldots, 612$ is shown in Figure 5.1.

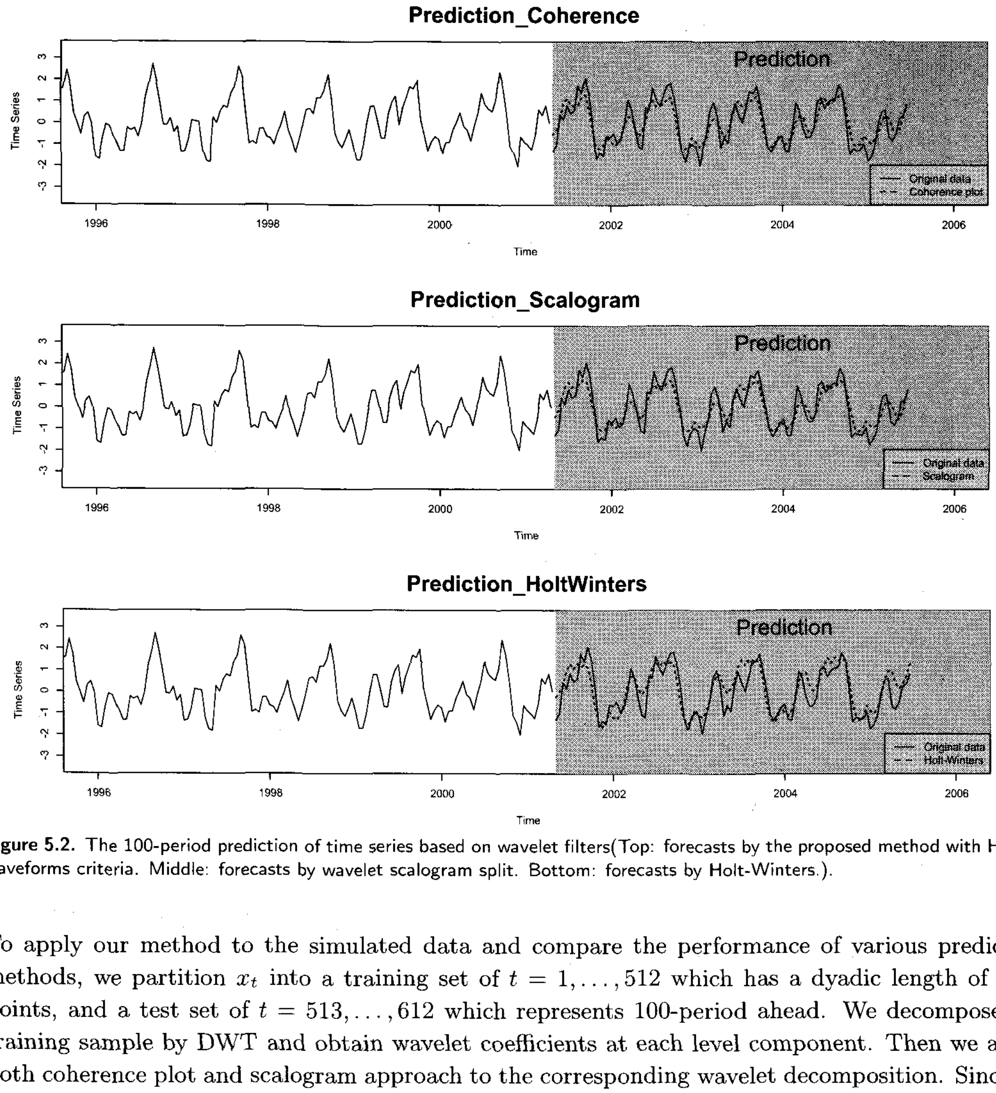

Figure 5.2. The 100-period prediction of time series based on wavelet filters(Top: forecasts by the proposed method with High-C waveforms criteria. Middle: forecasts by wavelet scalogram split. Bottom: forecasts by Holt-Winters.).

To apply our method to the simulated data and compare the performance of various prediction methods, we partition x_t into a training set of $t = 1, \ldots, 512$ which has a dyadic length of data points, and a test set of $t = 513, \ldots, 612$ which represents 100-period ahead. We decompose the training sample by DWT and obtain wavelet coefficients at each level component. Then we apply both coherence plot and scalogram approach to the corresponding wavelet decomposition. Since we intend the simulated time series to have no long-term trend, we only use reconstructed short-term seasonality components for forecasting. For the selected filter bank components, a bivariate VAR(1) model is imposed. Figure 5.2 shows the resulting 100-period forecasts, together with the training data and the actual observations in the lead period. We also present the forecasts of Holt-Winters for the same data in Figure 5.2. The PEs for each forecasting by the coherence plot, scalogram, and Holt-Winters are 0.22, 0.24 and 0.34, respectively. As can be seen, the coherence plot method yields the most accurate prediction in this case. The performance of Holt-Winters is primarily due to its heavy dependence on the last day's observations in the training data. The forecasts by the proposed method are more robust because they are based on more persistent patterns. This fact is more clear from the result of 100 times iteratively sampled data. The logarithms of PEs (in 100 samples) for each method are presented by boxplots in Figure 5.3, which indicates that the prediction by coherence plot approach gives more stable and accurate result.

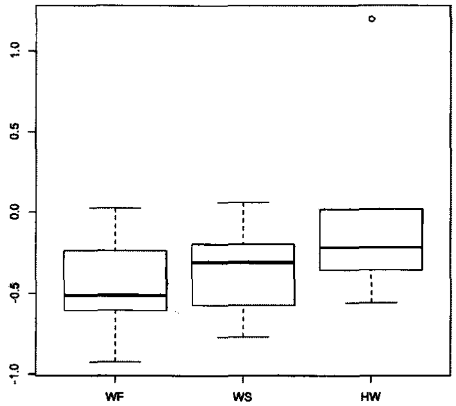

Figure 5.3. Boxplots for log(PE) of (wavelet filter) coherence plot approach(WF), wavelet scalogram approach(WS) and Holt-Winters(HW), respectively, in 100 samples.

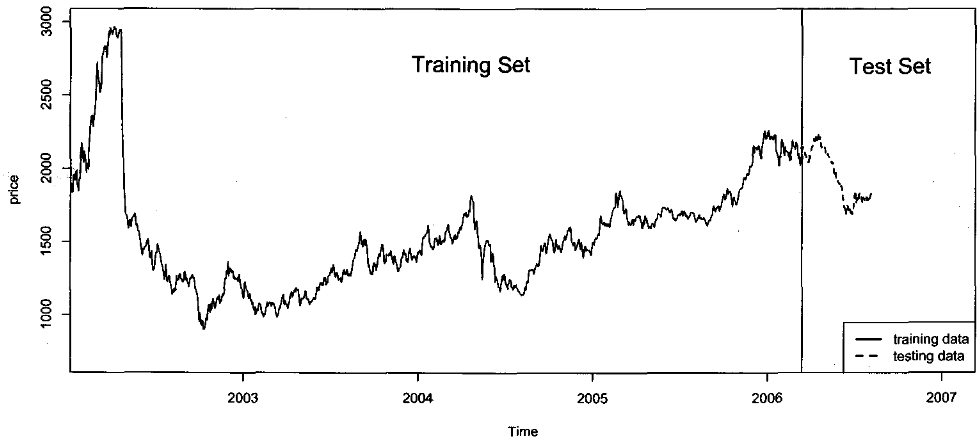

Figure 5.4. Daily stock price of consumer goods in Korea from the year 2002 to 2006.

5.2. Application to economic time series

We introduce a real data set to which the methodology proposed in this paper will be applied. In the previous section, we observed that the coherence plot approach outperforms the scalogram method though they all are based on the wavelet coefficients. So, we only consider the performance of the coherence plot approach applied to this real data here. The data set is a time series of the daily stock prices of consumer goods in Korea from January 2002 to August 2006, taken from the website of a stock price index provider called "FnGuide", `http://www.fnguide.com`. The plot of original data points with a sample size of $N = 1{,}124$ is presented in Figure 5.4. We represent the training sample of dyadic size 1,024 in a solid line and the test sample of size 100 in a dashed line. The prices of consumer goods soared around the year 2002, in which Korea co-hosted the 17^{th} World Cup with Japan.

After obtaining the wavelet decomposition $\boldsymbol{d} = \boldsymbol{W}\boldsymbol{y}$, we could get a coherence plot at each filter level as in Figure 5.5. As can be seen, five frequencies are statistically significant in χ^2-test. We

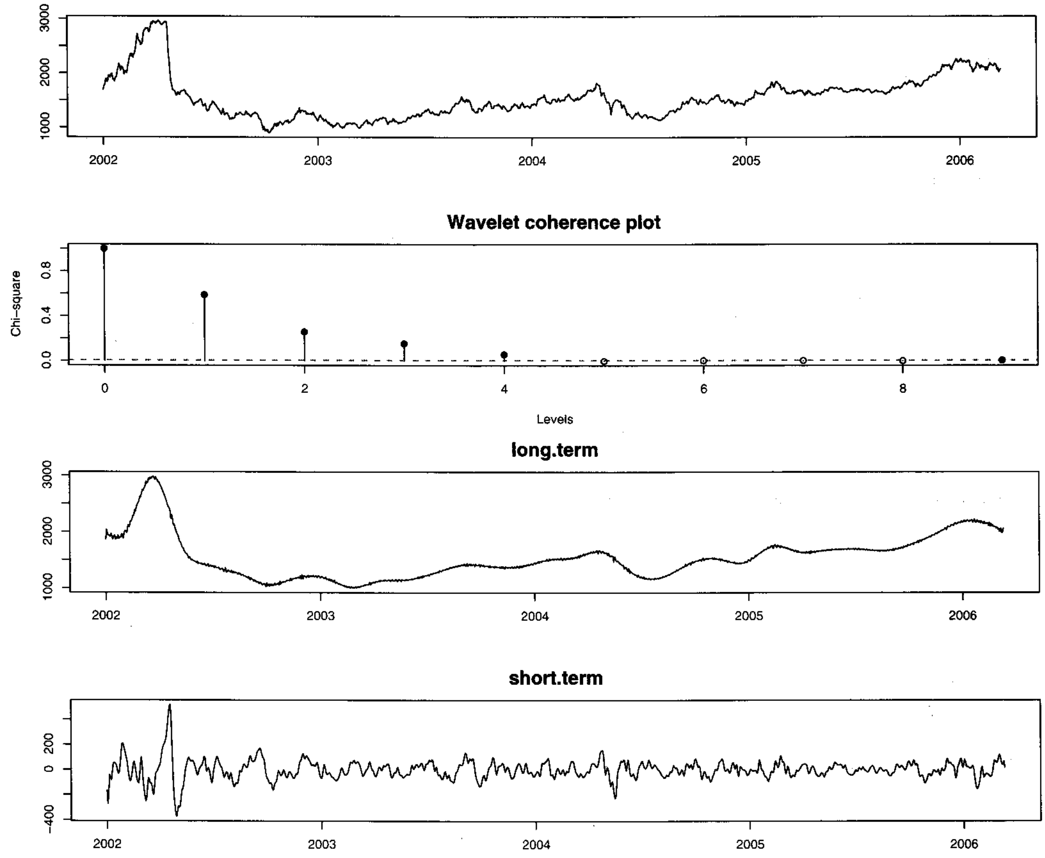

Figure 5.5. Decomposition of time series based on wavelet filter banks(Top: original time series. Second: coherence plot from High-C waveforms criteria, with a threshold in red dashed line. Bottom: the reconstructed split of long-term and short-term patterns.).

split the wavelet coefficients into two parts as described in Section 3 and removed the irregular term, and then applied the inverse DWT to $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ to obtain the reconstruction of long-term trend $\boldsymbol{y}^{(1)}$ and short-term seasonality $\boldsymbol{y}^{(2)}$ as follows:

$$\begin{aligned}\boldsymbol{y}^{(1)} &= W^{-1}\boldsymbol{d}^{(1)},\\ \boldsymbol{y}^{(2)} &= W^{-1}\boldsymbol{d}^{(2)}.\end{aligned}$$

Corresponding to these sequences, forecasts are modeled as a VAR(1) process with the optimal order $k = 1$ selected by the AIC criterion among $k = 0, 1, 2$. We use R package `vars` (Pfaff, 2008) for estimating parameters of VAR model and forecasting the decomposed time series. The forecasts are plotted with those obtained from Holt-Winters and SARIMA in Figure 5.6.

We also provide the prediction results of the proposed method with a comparison of those from Holt-Winters and SARIMA in terms of PE in Table 5.1. The prediction error of High-C waveforms criteria is the lowest 25736.61. This result implies that prediction by the proposed method outperforms the Holt-Winters and SARIMA model.

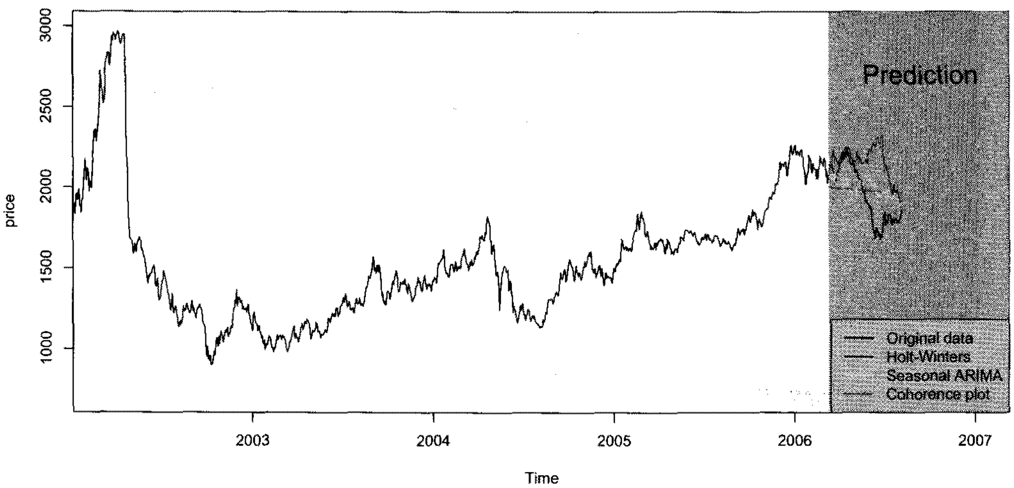

Figure 5.6. Prediction of time series based on each method.

Table 5.1. Prediction error of each forecasting method

	High-C waveforms	Holt-Winters	Seasonal ARIMA
Prediction Error(PE)	25736.61	83177.15	32468.87

6. Discussions

6.1. Grouping problems on coherence plot

In our method, the assignment of one coherence group to whether the long-term trend or short-term seasonality components is determined by the frequency structures of its waveform elements. Here, we observe the possibility that one can confuse which part the levels to be classified. For example, suppose that we obtain a drastically wiggly coherence plot. Since low- and high-frequencies are all mixed up in one coherence group in this situation, it is impossible to apply our method on that time series. However, we may choose subsets of each coherence group based on the principle of first preserving the High-C waveforms and then remove levels be remote from the majority of one group. This logic is based on the nested nature of multiresolution analysis framework where the wavelet decomposition is settled. However, those cases are rare and the sophisticated procedure for the treatment of those situations may not be simple, more studies need to be followed. This remains as a further research.

6.2. Boundary problems

It is well known that most nonparametric smoothing approaches including wavelet regression suffer from boundary problems or edge effects. The bias problems at the boundaries are due to the function we decompose is assumed to be periodic on it's interval of definition. For long-term forecasting of seasonal time series, therefore, it is crucial to overcome these boundary effects since they have a certain influence on long-lasting changes in a time series. In this paper, we just use DWT with the default boundary handling option of periodic function. There are some literatures introduced to handle the boundaries when using wavelet regression. Among others, the polynomial boundary treatment for wavelet regression proposed by Oh *et al.* (2001) can be considered to be a choice for this matter. We leave the research on the combination of their method and ours as a future work.

7. Concluding Remarks

In this paper, we present a method to extract meaningful patterns in a time series through the use of wavelet theory. This method consists of selecting and splitting the wavelet coefficients of the data set to analyze with respect to a specified basis. The decision of how to split these wavelet coefficients is made after an analysis of its scalogram or High-C waveforms criteria. For a successful application of the proposed approach, we analyze a simulation and a real stock price data. The forecasts by the proposed method are more robust because they are based on more persistent patterns. However, since there are no clear rationale for scalogram splitting approach, the result was not better than that of the coherence plot approach though the two are all based on wavelets. The method will be much improved if the distributional characteristics of scalogram is considered together in the forecasting of seasonal time series. It remains as a future research.

References

Ariño, M. A., Morettin, P. A. and Vidakovic, B. (2004). Wavelet scalogram and their applications in economic time series, *Brazilian Journal of Probability and Statistics*, **18**, 37–51.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd ed., Springer-Verlag, New York.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia.

Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage, *Biometrika*, **81**, 425–455.

Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press, Princeton.

Hannan, E. J. and Quinn, B. G. (1979). The determination of the order of an autoregression, *Journal of the Royal Statistical Society, Series B*, **41**, 190–195.

Lee, G.-H. (1998). Decomposition and forecasting of corporate bond yields using wavelet, *Economic Analysis: The Bank of Korea*, **4**, 1–23.

Li, T. H. and Hinich, M. J. (2002). A filter bank approach for modeling and forecasting seasonal patterns, *Technometrics*, **44**, 1–14.

Lütkepohl, H. (2007). *New Introduction to Multiple Time Series Analysis*, 2nd ed., Springer, Berlin.

McCoy, E. J. and Walden, A. T. (1996). Wavelet analysis and synthesis of stationary long-memory processes, *Journal of Computational Graphical Statistics*, **5**, 26–56.

Nason, G. P. and von Sachs, R. (1999). Wavelets in time series analysis, *Philosophical Transactions of the Royal Society, London A*, **357**, 2511–2526.

Nason, G. P. and Silverman, B. W. (1995). The stationary wavelet transform and some statistical applications, In *Wavelets and statistics* (A. Antoniadis and G. Oppenheim, ed.), Springer Lecture Notes in Statistics, No. 103, 281–300.

Oh, H.-S., Naveau, P. and Lee, G.-H. (2001). Polynomial boundary treatment for wavelet regression, *Biometrika*, **88**, 291–298.

Pfaff, B. (2008). VAR, SVAR and SVEC models: Implementation within R package vars, *Journal of Statistical Software*, **27**, 1–32.

Silverman, B. W. (2005). The `EbayesThresh` package: Empirical Bayes thresholding and related methods.

Vidakovic, B. (1999). *Statistical Modeling by Wavelets*, Wiley, New York.