Multiscale Time Series Analysis of the Taiwan Stock Market: A Wavelet-Based Approach

Chueh-Yung Tsao* **Chun-Ru Lin**

Department Of Business Administration
Chang Gung University

Abstract

In this paper we propose the multiscale time series model (MSTM), which is a hybrid model combining wavelet analysis with time series models. The basic advantage of MSTM is that, while investigating complex adaptive systems, such as economic systems or financial markets, the signals (prices) obtained from the system are an aggregation of many heterogeneous agents' behaviors, and consequently a global model cannot efficiently handle such signals. By using the MSTM, the signals (prices) can be examined locally and different patterns can be revealed under different scales.

Keywords: multiscale time series model, wavelet, signal decomposition, ARIMA

1. Introduction

In general, most financial studies use returns of assets, instead of prices, to examine financial issues. The main reason for using returns is that the return on an asset is easier for public investors to understand and provides a complete and scale-free summary of the investment opportunity. The other reason is that return series have more desirable statistical properties, e.g., a stationary property, and can be handled more easily than price series (Campbell, Lo, and MacKinlay, 1997).

However, it is difficult to perfectly capture a return series pattern using a linear statistical model. The explanatory capability of a regression model for a return series is generally fairly low where traditional econometric models, such as the autoregressive integrated moving average model (ARIMA), the generalized autoregressive conditional heteroskedasticity (GARCH) model, etc., are used. Therefore, many nonlinear econometric models are used in analyzing financial time series, such as the Markov switching model or the threshold model. Besides, various computational intelligence methods, such as artificial neural networks or genetic programming are also devoted to this area. For example, Kaastra and Boyd (1996) provide an overview of a step-by-step methodology to design a neural network for forecasting economic time series data. Although the explanatory capability efficiently rises, a common weakness of using these nonlinear models is that their structures are complex and their parameters are hard to interpret in financial and economic applications.

In this paper we propose the multiscale time series model (MSTM), which is a hybrid model combining wavelet analysis with time series models. We first obtain low- and high-frequency signals from the price data using wavelet transformations. Next we apply a time series model to each of the series in order to capture different patterns in each series. The basic advantage of MSTM is that when investigating complex adaptive systems, such as economic systems or financial markets, the signals (prices) obtained from the system are an aggregation of many heterogeneous agents' behaviors, and consequently a global model cannot efficiently handle such signals. By using the MSTM, the signals (prices) can be examined locally and different patterns can be revealed under different scales.

The remainder of the paper is organized as follows. The MSTM is introduced in Section 2, followed by the empirical analysis and the results in Section 3. Our conclusions are presented in Section 4.

2. Methodology

2.1 Wavelets

A discrete wavelet transform (DWT) can be regarded as a mapping from the signal domain to the wavelet coefficient domain. Let x be a vector with a length of power 2, i.e. $N = 2^J$, $J \in \aleph$. The N wavelet coefficients w are obtained using the transformation

$$w = Wx.$$

Here, W is an $N \times N$ orthonormal matrix that can be organized into J+1 submatrices

$$W = [W_1, W_2, \ldots, W_J, V_J]^T$$

* Corresponding author, email: cytsao@mail.cgu.edu.tw.

where W_j is an $N/2^j \times N$ matrix, $j = 1, 2, \ldots, J$, and V_J is a row vector of N elements. The W_j, $j = 1, 2, \ldots, J$, are composed of $N/2^j$ wavelet filters that can be used to capture changes in the original series on a scale of length $\lambda_j = 2^{j-1}$, and the V_J comprise a scaling filter that can be used to capture the average of the original series on a scale of length $2^J = 2\lambda_J$. The vector of wavelet coefficients w may then be partitioned accordingly into J+1 vectors,

$$w = [w_1, w_2, \ldots, w_J, v_J]^T$$

where w_j, $j = 1, \ldots, J$, and v_J are vectors of wavelet coefficients with length $N/2^j$, $j = 1, \ldots, J$, and $N/2^J$, respectively. The w_j, $j = 1, \ldots, J$, are associated with variations on a scale of length λ_j and the v_J associates with averages on a scale of length $2^J = N$. Therefore, the wavelet coefficients for the fine wavelet filters (small scale) capture high frequency information and those for coarse filters (large scale) capture low frequency information. The coefficients for the scaling filter represent the overall trend of the signal. Thus, by using the DWT we can convert the time series data from the time domain to the time-frequency domain.

An inverse discrete wavelet transform (IDWT), on the contrary, can be regarded as a mapping from the wavelet coefficient domain to the signal domain:

$$x = W^T w.$$

The IDWT provides a perfect reconstruction of the original series. A multi-resolution analysis (MRA) of the original series may now be defined via

$$x = \sum_{j=1}^{J+1} d_j$$

where $d_j = W_j^T w_j$, $j = 1, 2, \ldots, J$, be the jth level wavelet *detail* attributable to changes in x at scale λ_j, and the final wavelet detail $d_{J+1} = V_J^T v_J$ has all of its elements equal to the sample mean of the observations. Let $s_j = \sum_{k=j+1}^{J+1} d_k$, $j = 1, 2, \ldots, J$, be the jth level wavelet *smooth*. While the wavelet detail d_j is associated with variations in the data, s_j can be regarded as a cumulative sum of these variations and will become smoother as j increases. Similar, the jth level wavelet *rough* can be defined as $r_j = \sum_{k=1}^{j} d_k$, $j = 1, 2, \ldots, J$, which characterizes the remaining lower-scale details. Overall, the observation x can be presented as a linear combination of wavelet detail and wavelet rough. That is,

$$x = d_j + s_j$$

for $j = 1, 2, \ldots, J$.

In sum, by using the DWT, we can transform the original series into wavelet coefficients that measure the magnitude of the variations under a different scale and time and average level of the data. The IDWT then reconstructs the data by summing up several series. Consider a situation where

$$x = d_1 + s_1.$$

The original series is partitioned into two series: the high-frequency signal d_1 that captures the short-term behavior of the data at scale 1 and the low-frequency signal s_1 that captures the long-term behavior of the data at scales greater than 1.[1]

2.2 Autoregressive Integrated Moving-Average (ARIMA) Model

The autoregressive integrated moving-average (ARIMA) model, initiated by Box and Jenkins (1976), analyzes and forecasts univariate time series. The ARIMA(p, d, q) model for $\{x_t\}$ takes the form

$$r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i}$$

where $r_t \equiv (1-B)^d x_t$, B is the back-shift operator, $\{a_t\}$ is a white noise series, and p, q, and d are non-negative integers. ARIMA processes are found to be quite useful in modeling unit-root nonstationary time series data.[2]

An ARIMA model is said to exhibit unit-root nonstationarity because its AR polynomial has a unit root. A conventional approach for handling unit-root nonstationarity is to use differencing. Then, to make sure the series is stationary, we can inspect the series autocorrelation function (ACF) or apply the unit-root test. If we take AR (1) for example, the unit-root test involves considering the null hypothesis $H_0 : \phi_1 = 1$ versus the alternative hypothesis $H_a : \phi_1 < 1$; see Dickey and Fuller (1979).

The generalized autoregressive conditional heteroskedasticity (GARCH) models, proposed by Engle (1982) and Bollerslev (1986), are useful for modeling three stylized facts of financial time series, namely, the fat-tailed distribution of returns, the time-variant volatility of returns, and clustering outliers. Bollerslev, Chou, and Kroner (1992) found that the GARCH(1,1) is most identified in financial time series data. By combining the ARIMA and GARCH models, we can obtain the ARIMA(p, d, q)-GARCH(m, s) model. Its general form is

[1] Details and technicalities for wavelet analysis can be found, among other relevant books, in Percival and Walden (2000). Readers interested in the application in economics and finance can refer to Gencay, Selcuk, and Whitcher (2002).

[2] The Wold Representation Theorem states that any stationary series can be decomposed into the sum of a deterministic component and a stochastic component. The stochastic component can be represented as an MA(∞) process.

$$r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i}$$

$$a_t = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

where $\{\varepsilon_t\}$ is a iid random series with mean 0 and variance 1, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$.

We use Akaike information criterion (AIC) to determine the order p and q of an ARIMA process. AIC is defined as

$$\text{AIC} = \frac{-2}{T}\ln(\text{likelihood}) + \frac{2}{T} \times (\text{number of parameters})$$

where T is the sample size. We select the best model that has the minimum AIC value.

After fitting the selected model for each time series, we compare the fitting ability depending on the following three measures: mean square error (MSE), mean absolute deviation (MAD) and mean absolute percentage error (MAPE). These measures are defined as

$$MSE = \frac{1}{m}\sum_{j=1}^{m-1}(x_j - \hat{x}_j)^2$$

$$MAD = \frac{1}{m}\sum_{j=1}^{m-1}| x_j - \hat{x}_j |$$

$$MAPE = \frac{1}{m}\sum_{j=1}^{m-1}| \frac{\hat{x}_j}{x_j} - 1 |.$$

2.3 Multiscale Time Series Model (MSTM)

Generally speaking, implementing the multiscale time series model (MSTM) involves two steps. The first is to obtain low- and high-frequency signals from the original time series data by using DWT and IDWT. The two signals are also time series data and have the same observations as the original series. The second is to apply a time series model to each of the series in order to capture different patterns in each series.

The first step of the MSTM can be taken by choosing one wavelet filter among many, such as the Haar, Daubechies or least symmetric wavelet, combined with one kind of DWT, such as the basic DWT, the maximal overlap discrete wavelet transform (MODWT), or the discrete wavelet packet transform (DWPT). The second step of the MSTM is more flexible and the candidate models for this step involve huge time series models already existing in the literature, such as statistical models like the ARIMA or Markov switching models, or computational intelligence methods like neural networks or the symbolic regression model. In this analysis, we choose the most basic setting of the model, namely, the Haar wavelet combined with the basic DWT for the first step and the ARIMA model for the second step, in order to observe the preliminary performance of the MSTM. Other combinations could be considered in future research.

The idea of combining wavelet analysis with forecasting models is not new. Aussem and Murtagh (1997) and Aussem, Campbell, and Murtagh (1998) decompose the signal to be forecasted into its time scale components, and then use artificial neural networks on each component to produce an overall forecast. Kaboudan (2005) applies genetic programming and artificial neural networks to wavelet-transformed data. While all of the previous studies work on the wavelet coefficients, the MSTM works on the wavelet rough and wavelet detail, which is a more direct handling of the time series signals.

3. Empirical Analysis

3.1 Data

We examine the component stocks of the TSEC Taiwan 50 index in this study because the TSEC Taiwan 50 index is established by choosing the 50 highest market value stocks on the Taiwan Stock Exchange (TSE), which means that the component stocks are the most representative for the TSE. In order to fully collect 2,048 pieces of data daily from 1996/4/6 to 2003/12/23, only 24 stocks are considered in the experiment. The total sample for the experiment comprises 418,192 observations.[3]

3.2 Experiment Design

In the empirical analysis conducted in this study, we apply the Haar wavelet combined with the basic DWT to obtain low- and high-frequency signals. For each price series (x), we obtain

$$x = d_1 + s_1.$$

The high-frequency signal d_1 captures short-term behavior of the data on a scale of 1, i.e. the daily variations in the price series, and the low-frequency signal s_1 captures the long-term behavior of the data on a scale greater than 1, i.e. the aggregate variation in the price series on a scale of more than 1 day. Figure 1 is the time series plot of d_1 and s_1 for a sample stock. Notice that the high-frequency signal d_1 exhibits a stationary property and the low-frequency signal s_1 demonstrates nonstationary behavior. The unit-root test further approves this impression. The null hypothesis of an existing unit root is rejected for both

[3] Due to the size limitations of the paper, we do not provide details for the 24 stocks. Such detailed information is, however, available upon request from the authors.

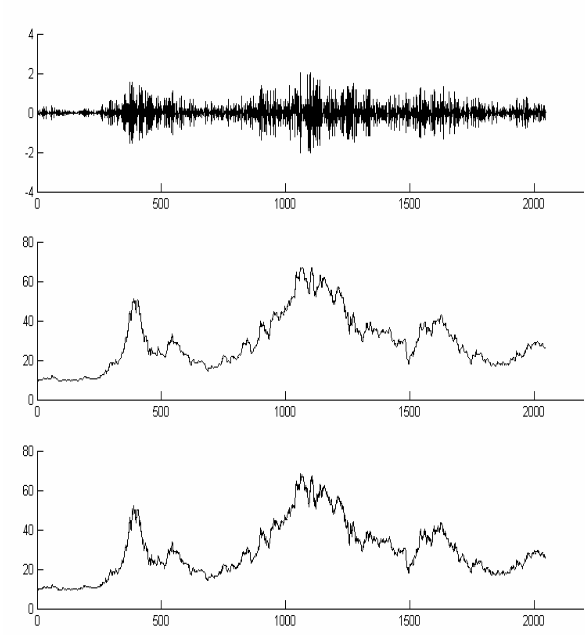

Fig. 1: Discrete Wavelet Transform of United Microelectronics Corp. (UMC).

d_1 and the first-order difference of s_1, which leads us to conclude that d_1 is a stationary process (d = 0) and that s_1 has one unit root (d = 1). Other stocks share similar properties. The ARIMA model is applied to both d_1 and s_1 to examine different time series patterns existing in each signal. The AIC rule is employed to determine the AR order p and MA order q.[4] In addition, the GARCH(1,1) model is considered in each series to model the time-variant volatility.

For comparison purposes, a standard econometric processing of the data is used as the benchmark model. What we do is first obtain continuous compounded returns from stock prices and then apply the ARIMA model combined with GARCH(1,1) to the return series. Usually, the return series is stationary and, hence, d = 0. In addition, the AIC rule is applied to determine the AR order p and MA order q. We refer to this benchmark model as the return model.

3.3 Feature Detection with MSTM

One advantage of the MSTM is that it is a local model, and so it can be expected that its fitting ability is greater than that of a global model if the data exhibits local characteristics. Table 1 presents the fitting results for both the MSTM and the return model. Three performance measures, namely, RMSE, MAD, and MAPE, are applied here. It should be noted that the MSTM outperforms the return model in all of the 24 stocks in terms of the three performance measures. The MSTM, on average, beats the return model by 18%, 33% and 33% for these three measures, respectively.

Another advantage of using the MSTM is that series characteristics, such as information transparency, complexity, etc., under different frequencies can easily be revealed. To investigate the complexity of the data, we simply use the order of the model as a measure of the degree of complexity. A higher order indicates more complexity and a lower order indicates less complexity of the data. In this study, the ARIMA model is applied and the degree of complexity can easily be obtained by adding the order of the AR term (p) and the order of the MA term (q). It is found, on average, that the order of the high-frequency series (3.54) is lower than the order of the low-frequency series (3.88). This finding suggests that the information existing in high-frequency series lasts for a shorter period of time and can easily be detected by using a simple linear model, and the information existing in low-frequency series lasts for a longer period of time and is harder to detect. The coefficients of determination (R^2) for the high- and low-frequency models further confirm this viewpoint. The average R^2 values for the high- and low-frequency models are about 0.34 and 0.05, respectively. For the benchmark model, i.e. the return model, it is interesting to find that its degree of complexity is just a mixture of the MSTM. The average order is 3.75, which is a little above of the mean order of the high- and low-frequency models (3.71). However, the fitting ability is fairly low (R^2 = 0.01) relative to the MSTM.

4. Conclusions

In this paper, a local model for financial time series is proposed. By using this model, different patterns existing in different frequency signals can easily be detected. Therefore, we refer to this as the multiscale time series model (MSTM). By choosing the Taiwan stock market as the reason for examining the performance of the MSTM, we find that the MSTM can fit the data well and that it can distinguish different complexities among high- and low-frequency series. The MSTM is so flexible that we can suitably choose each component of the model based on the data characteristics.

References

[1] A. Aussem, J. Campbell, and F. Murtagh, “Wavelet-based Feature Extraction and De-composition Strategies for Financial Forecasting,” *Journal of Computational Intelligence in Finance*, March/April, pp. 5-12, 1998.

[4] In this study, we let p and q range from 0 to 4. The AIC rule is applied to search for the model with the best fit among these 24 candidates.

Table 1: The fitting results of the MSTM and the return model.

	Return Model			MSTM		
Code	**RMSE**	**MAD**	**MAPE**	**RMSE**	**MAD**	**MAPE**
1216	0.022	0.016	0.016	0.017	0.009	0.009
1301	0.023	0.017	0.017	0.017	0.009	0.009
1303	0.024	0.018	0.018	0.019	0.010	0.010
1326	0.026	0.019	0.019	0.021	0.011	0.011
1402	0.030	0.023	0.023	0.024	0.013	0.013
1605	0.028	0.021	0.021	0.022	0.012	0.012
2201	0.025	0.019	0.019	0.019	0.011	0.011
2204	0.024	0.017	0.017	0.019	0.010	0.010
2301	0.033	0.026	0.026	0.027	0.015	0.015
2303	0.029	0.022	0.022	0.025	0.018	0.018
2308	0.031	0.024	0.024	0.026	0.019	0.019
2311	0.031	0.024	0.024	0.025	0.014	0.014
2317	0.029	0.022	0.022	0.025	0.015	0.015
2323	0.031	0.024	0.024	0.027	0.019	0.019
2324	0.030	0.023	0.023	0.026	0.018	0.018
2325	0.033	0.026	0.026	0.029	0.022	0.022
2330	0.028	0.021	0.021	0.024	0.015	0.015
2344	0.033	0.026	0.026	0.029	0.021	0.021
2371	0.026	0.019	0.019	0.020	0.011	0.011
2603	0.027	0.020	0.020	0.022	0.011	0.011
2609	0.026	0.020	0.020	0.021	0.011	0.011
2610	0.023	0.017	0.017	0.019	0.010	0.010
2801	0.026	0.019	0.019	0.021	0.012	0.012
9904	0.028	0.021	0.021	0.024	0.017	0.017
Ave.	0.028	0.021	0.021	0.023	0.014	0.014

[2] A. Aussem and F. Murtagh, "Combining Neural Network Forecasts on Wavelet-transformed Time Series," *Connection Science* 9, No. 1, pp. 113-121, 1997.

[3] T. Bollerslev, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics* 31, pp. 307-327, 1986.

[4] T. Bollerslev, R. Y. Chou, K. F. Kroner, "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics* 52, pp. 5-59, 1992.

[5] G. E. P. Box, G. M. Jenkins, "Time Series Analysis: Forecasting and Control," Holden-Day, San Francisco, 1976.

[6] J. Y. Campbell, A. W. Lo, and A. C. MacKinlay, *The Econometrics of Financial Markets*, Princeton University Press: New Jersey, 1997.

[7] C. Chen, R. E. McCulloch, and R. S. Tsay, "A Unified Approach to Estimating and Modeling Univariate Linear and Nonlinear Time Series," *Statistica Sinica* 7, pp. 451−472, 1997.

[8] Z. Ding, C. W. J. Granger, and R. F. Engle, "A Long Memory Property of Stock Returns and a New Model," *Journal of Empirical Finance* 1, pp. 83-106, 1993.

[9] D. A. Dickey and W. A. Fuller, "Distribution of the Estimates for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, pp. 427-431, 1979.

[10] R. F. Engle, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica* 50, pp. 987-1008, 1982.

[11] R. Gencay, F. Selcuk, and B. Whitcher, *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Academic Press, 2002.

[12] J. D. Hamilton, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica* 57, pp. 357-384, 1989.

[13] I. Kaastra and M. Boyd, "Designing a Neural Network for Forecasting Financial and Economic Time Series," 1996.

[14] M. Kaboudan, "Extended Daily Exchange Rates Forecasts Using Wavelet Temporal Resolutions," *New Mathematics and Natural Computation*, Vol. 1, No. 1, pp. 79-107, 2005.

[15] D. B. Percival, and A. T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge, 2000.