Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 35 (2008) 2080–2088
Expert Systems with Applications
www.elsevier.com/locate/eswa

Integrating GA-based time-scale feature extractions with SVMs for stock index forecasting

Shian-Chang Huang [a,*], Tung-Kuang Wu [b]

[a] Department of Business Administration, College of Management, National Changhua University of Education, No. 2, Shi-Da Road, Changhua 500, Taiwan
[b] Department of Information Management, National Changhua University of Education, Changhua, Taiwan

Abstract

By integrating genetic algorithm (GA)-based optimal time-scale feature extractions with support vector machines (SVM), this study develops a novel hybrid prediction model that operates for multiple time-scale resolutions and utilizes a flexible nonparametric regressor to predict future evolutions of various stock indices. The time series of explanatory variables are decomposed using wavelet bases, and a GA is employed to extract optimal time-scale feature subsets from decomposed features. These extracted time-scale feature subsets then serve as an input for an SVM model that performs final forecasting. Compared with neural networks, pure SVMs or traditional GARCH models, the proposed model performs best. The root-mean-squared forecasting errors are significantly reduced.

Keywords: Hybrid forecasting; Support vector machine; Wavelet analysis; Genetic algorithm, Time series forecasting

1. Introduction

Stock index predictions comprise challenging applications of modern time series forecasting and are essential to the success of many businesses and financial institutions. Numerous models have been developed that provide investors with increasingly precise forecasts. The objective of this study is to implement a novel forecasting strategy that fully exploits important features of international financial markets; international financial markets are highly correlated, and some markets are information leader of the others.

With the expansion of international financial relationships and continued liberalization of cross-border cash flow, increasingly strong correlations among financial markets help investors effectively forecast the co-movements of stock indices. However, global investors form a diverse group who operate on very different time scales. Consequently, the correlations between international market indices are not fixed over every time scale. The proposed forecasting model should extract these important time-scale features of international financial markets for a good prediction.

Wavelet analysis is a highly effective means of extracting time series features among various time scales. Wavelet analysis is extensively adopted in engineering applications, including signal processing and image compressions. However, wavelet analysis is relatively new to the fields of economics and finance. Ramsey and Zhang (1997) investigated foreign exchange data using waveform dictionaries. Davidson, Labys, and Lesourd (1998) analyzed the commodity price behavior using wavelet analysis. Ramsey and Lampart (1998a, 1998b) used wavelet analysis to decompose economic relationships of expenditure and income. Pan and Wang (1998) examined stock market inefficiency using wavelet analysis. Gençay, Selçuk, and Whitcher (2001, 2003, 2005) applied wavelet analysis to investigate scaling properties of foreign exchange volatility and systematic risk (the beta of an asset) in a capital asset pricing model. Recently, In and Kim (2006) and Kim and In (2003) utilized wavelet analysis to study the multiscale

* Corresponding author. Tel.: +886 4 7232105 7420, mobile: +886 953092968; fax: +886 4 7211290.
E-mail address: shhuang@cc.ncue.edu.tw (S.-C. Huang).

0957-4174/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2007.09.027

hedge ratio and relationship between financial variables and real economic activity. Lee (2004) employed wavelet analysis to study the transmission of stock market movements. Yamada (2005) used a wavelet-based beta estimation to investigate Japanese industrial stock prices.

For forecasting strategies, the Auto-Regressive Integrated Moving Average (ARIMA) technique has been widely used for time series forecasting. However, ARIMA is a general univariate model based on the assumption that forecasted time series are linear and stationary. In recent years, neural networks (NN) have been effectively applied in financial time series analysis and forecasting (Yao & Tan, 2000; Zhang & Hu, 1998; Zimmermann, Neuneier, & Grothmann, 2001). Tang and Fishwich (1993), Jhee and Lee (1993), Wang and Leu (1996), Hill, O'Connor, and Remus (1996), Kamruzzaman and Sarker (2003) and many other researchers have shown that NNs perform better than ARIMA models, specifically, for more irregular series and for multiple-period-ahead forecasting.

However, conventional NN models have several limitations: (I) dependency on a large number of parameters, e.g., network size, learning parameters and initial weights chosen, (II) possibility of being trapped into a local minima resulting in very slow convergence, and (III) over-fitting on training data resulting in a poor generalization ability. Recently, the support vector machine (SVM) (Cristianini & Shawe-Taylor, 2000; Schoelkopf, Burges, & Smola, 1999; Vapnik, 1998) has gained popularity and is regarded as a state-of-the-art technique for regression and classification applications. The formulation of an SVM embodies the structural risk minimization principle, thus combining excellent generalization properties with a sparse model representation. Tay and Cao (2001) used SVMs in forecasting financial time series. Their numerical results indicated that SVMs are superior to a multi-layer back-propagation NNs. Related works on time series forecasting include: Mohandes, Halawani, Rehman, and Hussain (2004), Pai and Lin (2005), Pai, Lin, Hong, and Chen (2006), Wang (2005).

The major innovation of this study lies in combing GA-based wavelet feature extractions with SVMs to develop a novel hybrid forecasting model. Wavelet analysis is used as a preprocessing to decompose and extract important time-scale features from the explanatory variables. However, given the high dimensionality of the new features, GA-based feature selection is utilized to prune irrelevant and noisy features, and produce effective feature subsets. In the empirical analysis, this study uses two data sets to test the new model. The first data set contains major Asian stock indices and the NASDAQ index. The second data set contains G7 indices.

The stock indices for each data set are strongly correlated. Some indices are information leaders for the others. For forecasting a specific index, wavelet analysis is applied to decompose and extract the important features from all lagged returns in the data set, and numerous time-scale features are selected using a GA to form new regressors with improved prediction power. Finally, selected feature subsets are fed into an SVM model for nonparametric forecasting. Compared with NN systems, the SVM model minimizes structural risk in the prediction model, as opposed to minimizing empirical risk typical of NN systems. Thus the novel forecasting model eliminates the over-fitting problem associated with traditional nonparametric regression models.

The rest of the paper is organized as follows. Section 2 describes GA-based time-scale feature extraction and the SVM model. Section 3 then introduces traditional GARCH prediction models in finance. Next, Section 4 describes the data used in the study, and discusses empirical findings. Conclusions are finally drawn in Section 5.

2. GA-based time-scale feature extraction

We first introduce a basic tool of wavelet analysis: the multiresolution decomposition (MRD). For a thorough review of wavelet analysis we refer to Chui (1992), Daubechies (1992), and Percival and Walden (2000). Practical applications of wavelet analysis is given in Lee (1998) and Gençay et al. (2002). Technical details of wavelet analysis are discussed in Bruce and Gao (1996).

2.1. Multiresolution decomposition

Any function $f(t)$ in $L^2(R)$ can be decomposed by a sequence of projections onto the wavelet basis. The wavelet representation of the signal or function $f(t)$ in $L^2(R)$ can be written as

$$f(t) = \sum_k s_{J,k}\phi_{J,k}(t) + \sum_k d_{J,k}\psi_{J,k}(t) + \sum_k d_{J-1,k}\psi_{J-1,k}(t) + \cdots + \sum_k d_{1,k}\psi_{1,k}(t),$$

where ϕ is the father wavelet and ψ the mother wavelet. $\phi_{j,k}$ and $\psi_{j,k}$ are scaling and translation of ϕ and ψ, defined as

$$\phi_{j,k}(t) = 2^{-j/2}\phi(2^{-j}t - k) = 2^{-j/2}\phi\left(\frac{t - 2^j k}{2^j}\right), \tag{1}$$

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k) = 2^{-j/2}\psi\left(\frac{t - 2^j k}{2^j}\right). \tag{2}$$

In the representation J is the number of multiresolution components, and $s_{J,k}$ are called the smooth coefficients, and $d_{j,k}$ are called the detailed coefficients. If we define

$$S_J(t) = \sum_k s_{J,k}\phi_{J,k}(t), \tag{3}$$

$$D_j(t) = \sum_k d_{j,k}\psi_{j,k}(t) \quad \text{for } j = 1, 2, \ldots, J. \tag{4}$$

The functions (3) and (4) are called the smooth signal and the detail signals, respectively, which constitute a decomposition of a signal into orthogonal components at different scales. A signal $f(t)$ can thus be expressed in terms of these signals:

$$f(t) = S_J(t) + D_J(t) + D_{J-1}(t) + \cdots + D_1(t). \tag{5}$$

2.2. Optimal feature subset selection by genetic algorithm

In general, feature (variable) selection is an important aspect of regression problems, since the features selected are used to build the regressors. Careful consideration should be given to the problem of feature subset selection with high-dimensional data. On the other hand, whereas two variables could be considered good predictors individually, there could be little to gain by combining the two variables together in a feature vector. Especially, when these variables are highly multicollinear. In our high-dimensional features, some wavelet features are indeed highly correlated. We need a GA to prune irrelevant and noisy features as well as producing effective feature subsets.

Let's denote D_j and S_J as the new feature vectors. In our case, after wavelet decomposition, there are $n \times (J+1)$ (n is the total number of indices used as regressors) feature vectors, GA selects only subset of vectors from this $n \times (J+1)$-dimensional space, and then the selected subset form the new regressors. There are $2^{n\times(J+1)}$ (it is more than 10,000 if 20-D space is searched, and 3×10^{10} for 35-D feature space) candidates to form the new regressors. Obviously, it is not feasible to evaluate all combinations to find the optimum subset for index prediction. GA is a good option to sort this problem out.

There are many forecasting performance indices can be used as fitness function, such as MSE (mean-squared error), RMSE (root-mean-squared error), MAE (mean-absolute error), and MAPE (mean-absolute percent error). In this article, RMSE is employed as the performance measure. It is defined as

$$\mathrm{RMSE} = \left(\frac{1}{N} \sum_{t=1}^{N} (r_t - \hat{r}_t)^2 \right)^{1/2}, \tag{6}$$

where N is the number of forecasting periods, r_i is the actual return at period t, and $\hat{r}_t$ is the forecasting return at period t. In this study, we adopted the following fitness function

$$\mathrm{Fitness} = \frac{1}{\mathrm{RMSE}}. \tag{7}$$

2.3. Support vector machines

The support vector machines (SVMs) were proposed by Vapnik (1998). Based on the structured risk minimization (SRM) principle, SVMs seek to minimize an upper bound of the generalization error instead of the empirical error as in other neural networks. Additionally, the SVMs models generate the regress function by applying a set of high dimensional linear functions. The SVM regression function is formulated as follows

$$y = w\phi(x) + b, \tag{8}$$

where $\phi(x)$ is called the feature, which is nonlinear mapped from the input space x to the future space. The coefficients w and b are estimated by minimizing

$$R(C) = C\frac{1}{N}\sum_{i=1}^{N} L_\varepsilon(d_i, y_i) + \frac{1}{2}\|w\|^2, \tag{9}$$

where

$$L_\varepsilon(d, y) = \begin{cases} |d-y| - \varepsilon, & |d-y| \geqslant \varepsilon, \\ 0, & \text{others}, \end{cases} \tag{10}$$

where both C and ε are prescribed parameters. The first term $L_\varepsilon(d,y)$ is called the ε-intensive loss function. The d_i is the actual price in the ith period. This function indicates that errors below ε are not penalized. The term $\frac{C}{N}\sum_{i=1}^{N} L_\varepsilon(d_i, y_i)$ is the empirical error. The second term, $\frac{1}{2}\|w\|^2$, measures the smoothness of the function. C evaluates the trade-off between the empirical risk and the smoothness of the model. Introducing the positive slack variables ξ and ξ^*, which represent the distance from the actual values to the corresponding boundary values of ε-tube. Eq. (9) is transformed to the following constrained formation

$$\min_{w,b,\xi,\xi^*} R(w, \xi, \xi^*) = \frac{1}{2}w^{\mathrm{T}}w + C\left(\sum_{i=1}^{N}(\xi_i + \xi_i^*)\right). \tag{11}$$

Subject to

$$w\phi(x_i) + b_i - d_i \leqslant \varepsilon + \xi_i^*, \tag{12}$$

$$d_i - w\phi(x_i) - b_i \leqslant \varepsilon + \xi_i, \tag{13}$$

$$\xi_i, \xi_i^* \geqslant 0. \tag{14}$$

After taking the Lagrangian and conditions for optimality, One can get the dual representation of the model

$$y = f(x, \alpha, \alpha^*) = \sum\nolimits_{i=1}^{N} (\alpha_i - \alpha_i^*) K(x, x_i) + b, \tag{15}$$

where α_i and α_i^* are Lagrangian multipliers, which are the solution to the dual problem, and $K(x,x_i)$ is the kernel function. b follows from the complementarity Karush–Kuhn–Tucker (KKT) conditions.

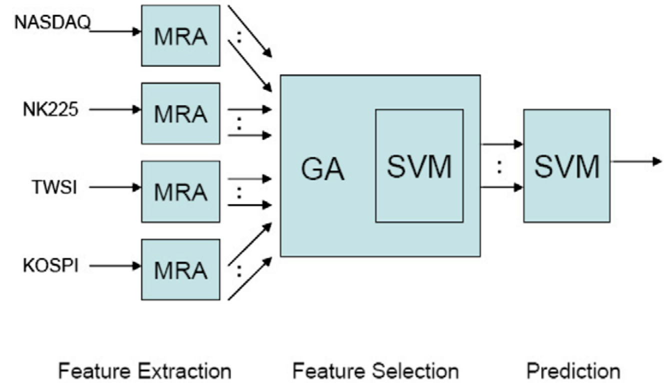

Fig. 1. Flow diagram of the hybrid model.

The value of the kernel is equal to the inner product of two vectors x_i and x_j in the feature space, such that $K(x_i,x_j) = \phi(x_i)\phi(x_j)$. Any function that satisfying Mercer's condition (Vapnik, 1998) can be used as the Kernel function. The Gaussian kernel function

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{16}$$

is specified in this study, because Gaussian kernels tend to give good performance under general smoothness assumptions.

The flow diagram of the proposed model is displayed in Fig. 1.

3. GARCH prediction models form finance

For comparison, a generalized autoregressive conditional heteroscedasticity models (Bollerslev, 1986) from finance are also used for predictions. The following model is employed to predict daily returns of an index. In the conditional mean part

$$r_t = \alpha + \beta r_{t-1} + \gamma x_{t-1} + \varepsilon_t, \tag{17}$$

this equation describes the causal relationship between current returns and lagged returns among market indices. r_t is the daily return at time t for each index, x_{t-1} is the lagged NASDAQ or lagged S&P 500 return which serve as an another explanatory variable, and $\varepsilon_t \sim N(0,\sigma_t)$, the innovation or shock at time t. σ_t follows a univariate GJR–GARCH(1,1) process

$$\sigma_t = a_0 + a_1\varepsilon_{t-1}^2 + c_1 S_{t-1}\varepsilon_{t-1}^2 + b_1\sigma_{t-1}^2, \tag{18}$$

where

$$S_{t-1} = \begin{cases} 1 & \text{if } \varepsilon_{t-1} < 0, \\ 0 & \text{if } \varepsilon_{t-1} \geqslant 0. \end{cases} \tag{19}$$

That is, depending on whether ε_{t-1} is above or below the threshold value of zero, ε_{t-1}^2 has different effects on the conditional variance σ_t^2. The asymmetric impact of ε_{t-1} on σ_t^2 has known as the leverage effects.

4. Experimental results and analysis

Two data set are used in this study. The first data set comprises the major Asian stock indices, including the following daily indices: NASDAQ (US), NK225 (Japan), TWSI (Taiwan) and KOSPI (South Korea). All index data are for the period from January 2003 to December 2004. There are 435 observations. These stock market indices are then transformed into daily returns. Figs. 2 and 3 plot the returns series. Table 1 presents the inter-correlations between the indices.

The second data set used in this study comprises G7 stock indices including the following daily indices: CAC40 (France), FTSE100 (UK), DAX30 (Germany), MIB40 (Italy), TSX60 (Canada), S&P500 (US), and NK225 (Japan). All index data is for period from January 2004 to December 2005. This data set also has 435 observations. Table 2 shows the inter-correlations between the indices.

This study considers one-step-ahead forecasting, because one-step-ahead forecasting can prevent problems associated with cumulative errors from the previous period for out-of-sample forecasting. The SVM model is trained in a batch manner, namely, 300 data points prior to the day of prediction are treated as the training data set. The window of training data set slides with the current prediction. Daily returns for the last 135 days are used as the test data set for evaluating the performances of the prediction models.

This study applies the Daubechies least asymmetric filters with length 8 to decompose explanatory variables. In each data set, the lagged returns of every index, (r_{t-1}), serve as the explanatory variables for prediction. These returns are then decomposed into five mutually orthogonal

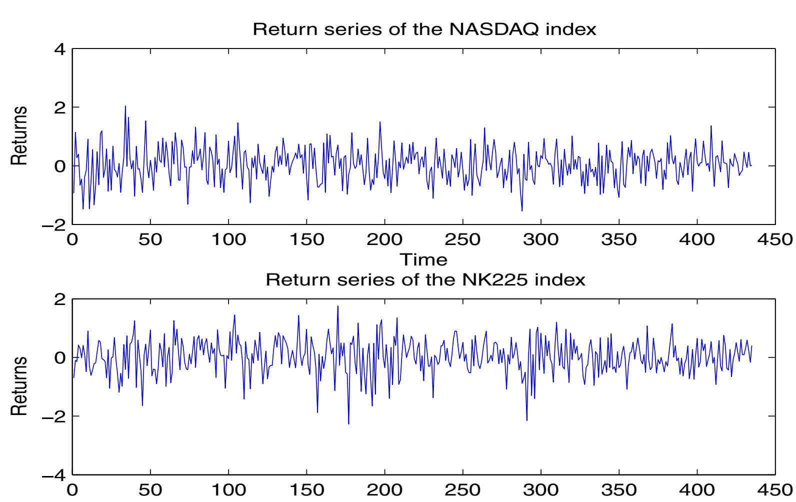

Fig. 2. Return series of NASDAQ and NK225 indices.

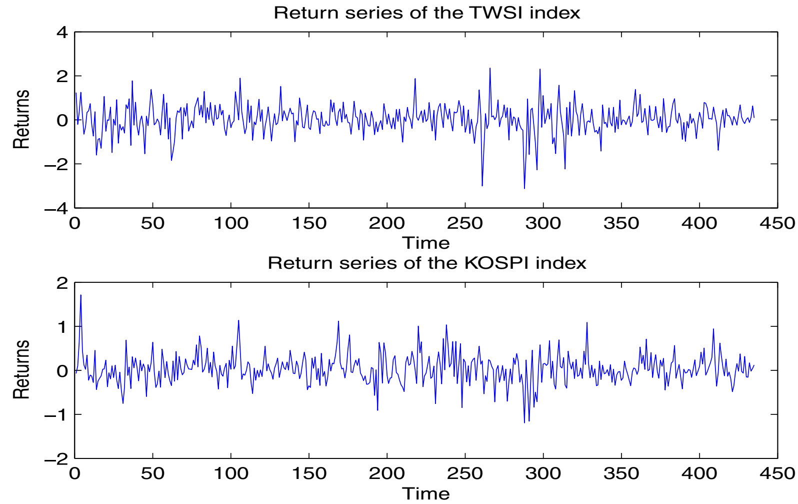

Fig. 3. Return series of TWSI and KOSPI indices.

Table 1
Correlations on raw data (major Asian indices)

	NASDAQ	NK 225	TWSI	KOSPI
NASDAQ	1.0000	0.2195	0.2807	0.1432
NK 225	0.2195	1.0000	0.4951	0.3852
TWSI	0.2807	0.4951	1.0000	0.3465
KOSPI	0.1432	0.3852	0.3465	1.0000

periodicity series, ranging from the shortest-periodicity series to the longest-periodicity series. Due to decomposition, the proposed method captures important time-scale features that cannot be revealed in aggregated explanatory variables. These features are then selected by a wrapper GA to form the optimal feature subset. Fig. 1 presents the detailed flow diagram of the hybrid model.

To model the nonlinear relationships between these features and output data, an SVM model is applied to forecast the future evolution of a target stock index. The optimal feature subset extracted by GA serves as SVM model inputs. The parameters utilized in GA are set as follows: population size = 200, chromosome mutation chance = 1/(size+1) (where size is the number of features in the chromosome), the number of iterations = 100. The SVM parameters are selected by grid searching which are $\varepsilon = 0.01$, $C = 100$ and $\sigma = 10$ for the Gaussian Kernel. For comparison of model performance, this study adopts the RMSE as the performance measure.

Tables 3 and 4 summarize the forecasting performance of the GARCH model, the pure SVM model (using raw data as regressors and no wavelet-based feature extraction), the NN model (with radial basis functions), and the proposed hybrid model using both data sets. Figs. 4–6 present actual returns, predicted values and model residuals for the major Asian stock for the proposed model. Figs. 7–9 present similar results for the GARCH model for comparison.

For the first data set, the performance of the pure SVM is slightly worse that using the GARCH model. The performance of the NN model is unstable (especially for the TWSI index), and worse than the pure SVM model. The proposed hybrid model with a GA-based time-scale feature extraction performed best and significantly reduces the root-mean-squared forecasting errors. The performance ranking is similar for the second data set. The performances of the pure SVM and the NN are worse than that of the GARCH model in almost every index. The proposed hybrid model also performed the best reducing the forecasting errors by 20–30%.

Considering the first data set, the success of the proposed forecasting model can be attributed to the following reasons. First, the correlations between the NASDAQ and these Asian indices are complex. Wavelet analysis provides a solid foundation for extracting the important features between these indices. For the finest time scale, $D1$,

Table 2
Correlations on raw data (G7 indices)

	CAC 40	FTSE 100	DAX 30	MIB 40	TSX 60	S&P 500	NK 225
CAC 40	1.0000	0.7612	0.9265	0.8650	0.3188	0.3559	0.3546
FTSE 100	0.7612	1.0000	0.7251	0.7126	0.3039	0.2423	0.3313
DAX 30	0.9265	0.7251	1.0000	0.8519	0.3037	0.3725	0.3376
MIB 40	0.8650	0.7126	0.8519	1.0000	0.3475	0.3635	0.3176
TSX 60	0.3188	0.3039	0.3037	0.3475	1.0000	0.5723	0.1850
S&P 500	0.3559	0.2423	0.3725	0.3635	0.5723	1.0000	0.0993
NK 225	0.3546	0.3313	0.3376	0.3176	0.1850	0.0993	1.0000

Table 3
Relative forecasting performance of four models on major Asian stock indices

	NK 225	TWSI	KOSPI
GARCH	0.4725	0.5557	0.2673
NN	0.4604	0.7600	0.2701
Pure SVM	0.4426	0.5672	0.2883
GA SVM	0.2658	0.3870	0.1939

Table 4
Relative forecasting performance of three models on G7 indices

	CAC 40	FTSE 100	DAX 30	MIB 40	TSX 60	NK 225
GARCH	0.2970	0.2853	0.3335	0.2541	0.2582	0.3944
NN	0.3206	0.2963	0.3728	0.2827	0.3264	0.3825
Pure SVM	0.3213	0.3023	0.3708	0.2788	0.3394	0.3790
GA SVM	0.2516	0.2306	0.2544	0.1808	0.2341	0.2768

intraday traders, market makers or hedging strategists view NASDAQ returns as a leading signal to trade Asian stocks. According to Kim and In (2003), intraday traders and market makers look to both purchase and sell financial assets to realize quick profit (or minimize loss) over very short time scales ranging from seconds to hours. Similarly, hedging fund strategists also trade very often. Thus, the $D1$ components are usually strong mean-reverting, revealing their short-periodic behavior. On this time scale, the NASDAQ index is strongly correlated with these Asian indices, and leads these indices. Thus, NASDAQ index has substantial predictive power for these indices.

For intermediate time scales $D2$, $D3$ and $D4$, the main traders are international portfolio managers. Their trades typically occur on a weekly to monthly basis, with little attention paid to intra-day prices. Thus, correlation patterns are totally different from the finest and coarsest time scales. That is, the correlation patterns are not fixed over every time scale. Table 5 reveals this phenomenon clearly.

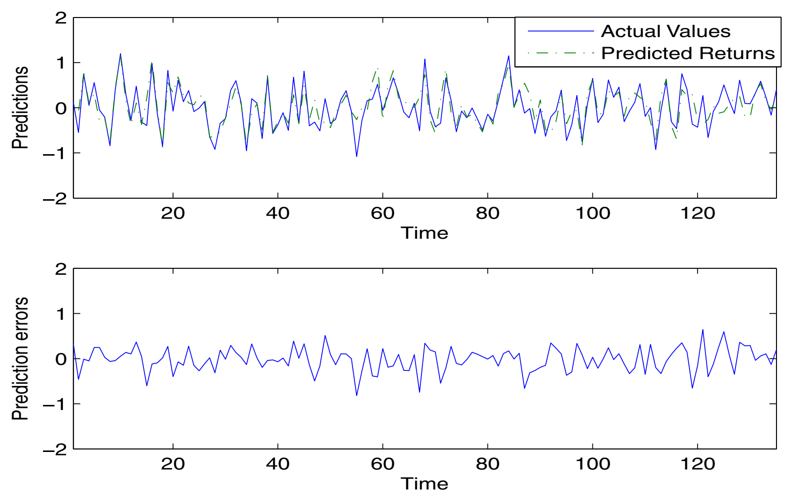

Fig. 4. Hybrid model forecasts on the NK225 index returns.

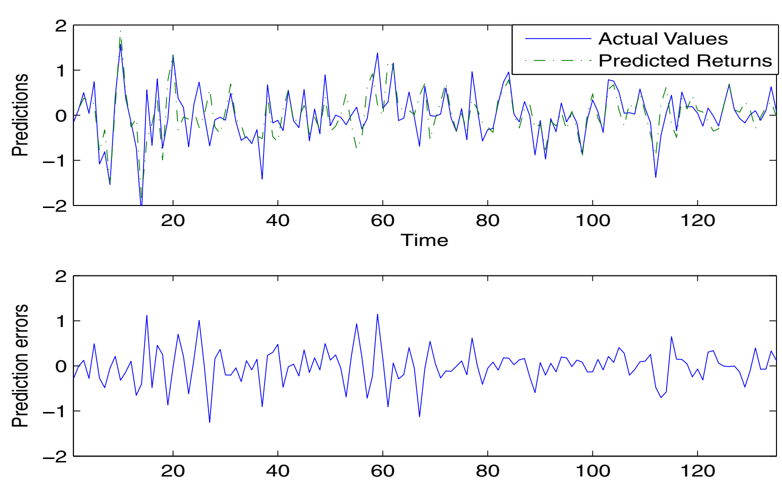

Fig. 5. Hybrid model forecasts on the TWSI index returns.

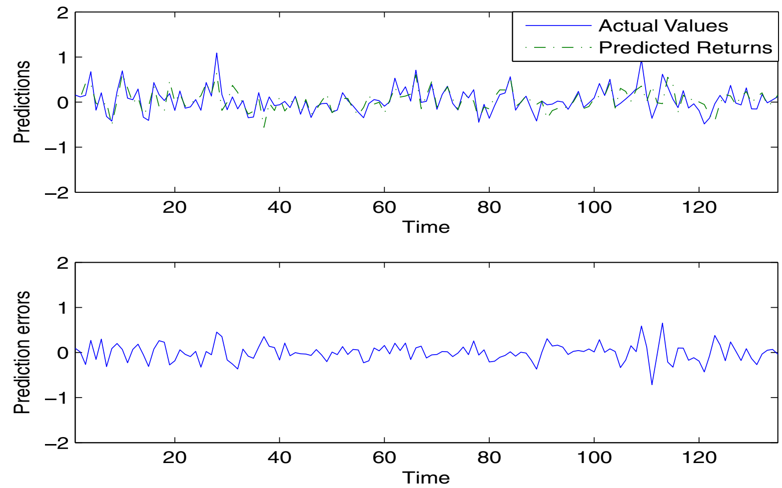

Fig. 6. Hybrid model forecasts on the KOSPI index returns.

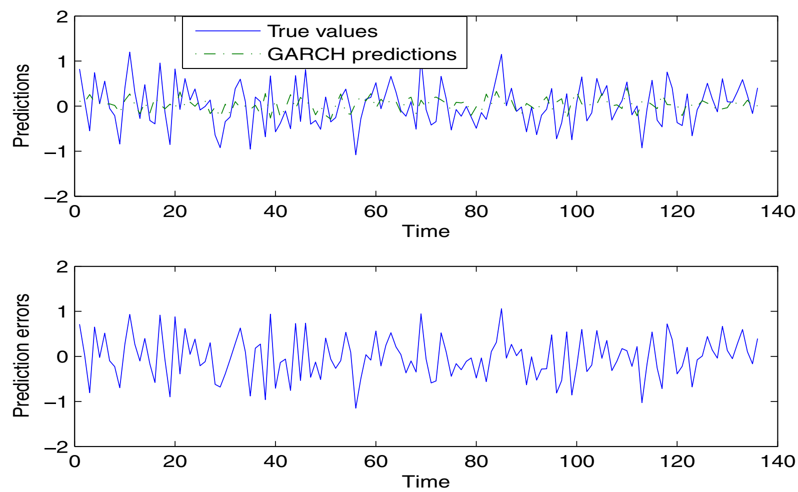

Fig. 7. GARCH forecasts on the NK225 index returns.

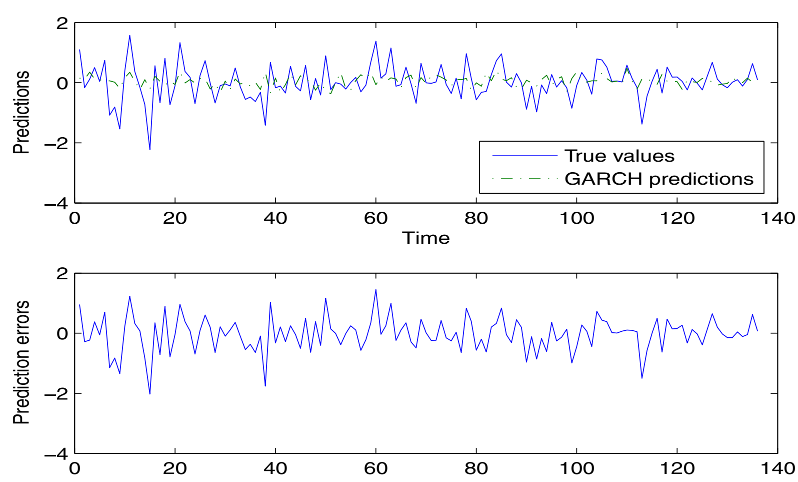

Fig. 8. GARCH forecasts on the TWSI index returns.

For the coarsest time scale $S4$, the principal traders are central banks operating on long time scales and often consider long-term economic fundamentals with long investment horizon. Since most high-tech investments

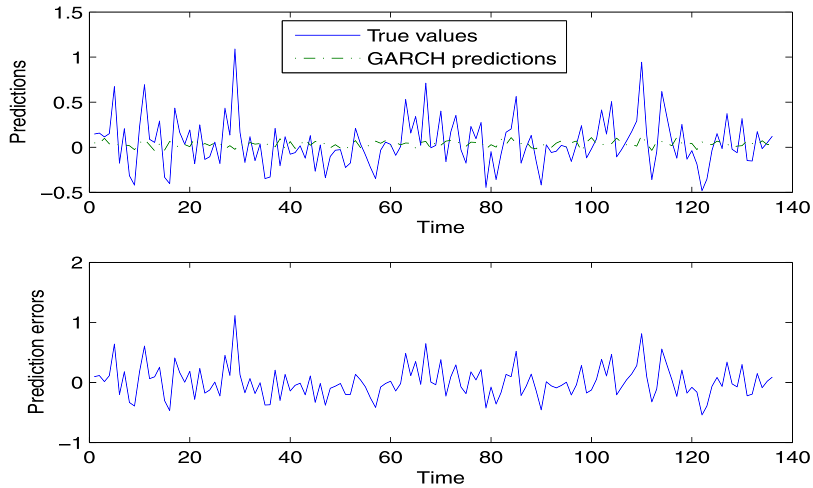

Fig. 9. GARCH forecasts on the KOSPI index returns.

worldwide have moved closer and closer, especially in Silicon Valley and Eastern Asia, the long-term correlations between these indices are also strong.

In summary, the correlation patterns of each data set are not fixed over every time scale. Combing with a GA, the proposed hybrid model captures the optimal feature subset and generates accurate forecasts. The conventional GARCH and pure SVM predictions only incorporate information about the "average" correlations among indices, and, thus, their performance is poor.

Second reason accounting for the excellent performance of the proposed model is that SVMs map input data into a high-dimensional reproducing kernel Hilbert space (RKHS) where a linear regression is performed. The RKHS has rich algebraic and topological structures for capturing nonlinear relationships between input and output data. Consequently, SVMs provide a valuable and flexible framework for representing relations among data.

Table 5
Correlations for different time scales between NASDAQ and every Asian index

	*a*4	*d*4	*d*3	*d*2	*d*1
NK225 (*a*4)	0.48	0.00	0.00	0.00	0.00
NK225 (*d*4)	0.00	0.77	0.00	0.00	0.00
NK225 (*d*3)	0.00	0.00	0.56	0.01	0.00
NK225 (*d*2)	0.00	0.00	0.00	0.37	0.00
NK225 (*d*1)	0.00	0.00	0.00	0.00	−0.05
TWSI (*a*4)	0.50	0.01	0.01	−0.01	0.00
TWSI (*d*4)	0.01	0.54	0.01	0.00	0.00
TWSI (*d*3)	0.00	0.00	0.57	0.00	0.00
TWSI (*d*2)	0.00	0.00	0.00	0.33	0.00
TWSI (*d*1)	0.00	0.00	0.00	0.00	0.10
KOSPI (*a*4)	0.27	0.00	0.02	−0.01	0.00
KOSPI (*d*4)	0.01	0.46	0.01	0.00	0.00
KOSPI (*d*3)	−0.01	0.01	0.36	0.03	0.00
KOSPI (*d*2)	0.00	−0.01	0.01	0.22	0.00
KOSPI (*d*1)	0.00	0.00	0.00	0.00	−0.03

5. Conclusions

Integrating GA-based time-scale feature extractions and SVMs, this study develops a novel hybrid prediction model that operates on multiple time-scale resolutions and utilizes a flexible nonparametric regression model to predict the future evolution of stock indices.

Global investors are a diverse group that operates on very different time scales. Thus, the correlation patterns among indices changes over each time scales. The GA is a effective mean for selecting the optimal wavelet features for SVM predictions. The GARCH and pure SVM predictions only incorporate information regarding the "average" correlations among the indices. Therefore, their performance is poor. By using wavelet analysis and by mapping data to the high-dimensional reproducing kernel Hilbert space, the proposed hybrid model captures all important features, and is excellent at forecasting.

In sum, the highly effective forecasting framework in this study can also be applied to other problems involving financial forecasting. Results of this study can also be used to perform a good hedge on global investments.

References

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*, 307–327.

Bruce, A., & Gao, H. Y. (1996). *Applied wavelet analysis with SPLUS*. New York: Springer-Verlag.

Chui, C. K. (1992). *An introduction to wavelets*. Boston, MA: Academic Press.

Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines*. Cambridge University Press.

Daubechies, I. (1992). *Ten lectures on wavelets*. Philadelphia, PA: SIAM.

Davidson, R., Labys, W. C., & Lesourd, J. B. (1998). Wavelet analysis of commodity price behavior. *Computational Economics, 11*, 103–128.

Gençay, R., Selçuk, F., & Whitcher, B. (2001). Scaling properties of foreign exchange volatility. *Physica, 289*, 249–266.

Gençay, R., Selçuk, F., & Whitcher, B. (2002). *An introduction to wavelets and other filtering methods in finance and economics*. London: Academic Press.

Gençay, R., Selçuk, F., & Whitcher, B. (2003). Systematic risk and time scales. *Quantitative Finance, 3*, 108–116.

Gençay, R., Selçuk, F., & Whitcher, B. (2005). Multiscale systematic risk. *Journal of International Money and Finance, 24*, 55–70.

Hill, T., O'Connor, M., & Remus, W. (1996). Neural network models for time series forecasts. *Management Science, 42*, 1082–1092.

In, F., & Kim, S. (2006). The hedge ratio and the empirical relationship between the stock and futures markets: A new approach using wavelet analysis. *Journal of Business, 79*, 799–820.

Jhee, W. C., & Lee, J. K. (1993). Performance of neural networks in managerial forecasting. *Intelligent Systems in Accounting, Finance and Management, 2*, 55–71.

Kamruzzaman, J., & Sarker, R. (2003). Forecasting of currency exchange rates using ANN: A case study. In *Proceedings of the IEEE International Conference on Neural Networks & Signal Processing (ICNNSP03)* (pp. 793–797). Nanjing, China.

Kim, S., & In, F. (2003). The relationship between financial variables and real economic activity: Evidence from spectral and wavelet analyses. *Studies in Nonlinear Dynamics and Econometrics 7*, Article 4.

Lee, G. H. (1998). Wavelets and wavelet estimation: A review. *Journal of Economic Theory and Econometrics, 4*, 123–158.

Lee, S. H. (2004). International transmission of stock market movements: A wavelet analysis. *Applied Economics Letters, 11*, 197–201.

Mohandes, M. A., Halawani, T. O., Rehman, S., & Hussain, A. A. (2004). Support vector machines for wind speed prediction. *Renewable Energy, 29*, 939–947.

Pai, P. F., & Lin, C. S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. *International Journal of Management Science, 33*(6), 497–505.

Pai, P. F., Lin, C. S., Hong, W. C., & Chen, C. T. (2006). A hybrid support vector machine regression for exchange rate prediction. *International Journal of Information and Management Sciences, 17*(2), 19–32.

Pan, Z., & Wang, X. (1998). A stochastic nonlinear regression estimator using wavelets. *Computational Economics, 11*, 89–102.

Percival, D., & Walden, A. (2000). *Wavelet methods for time series analysis*. Cambridge: Cambridge University Press.

Ramsey, J. B., & Lampart, C. (1998a). Decomposition of economic relationships by timescale using wavelets. *Macroeconomic Dynamics, 2*, 49–71.

Ramsey, J. B., & Lampart, C. (1998b). The decomposition of economic relationships by time scale using wavelets: Expenditure and income. *Studies in Nonlinear Dynamics and Econometrics, 3*, 23–42.

Ramsey, J. B., & Zhang, Z. (1997). The analysis of foreign exchange data using waveform dictionaries. *Journal of Empirical Finance, 4*, 341–372.

Schoelkopf, B., Burges, C. J. C., & Smola, A. J. (1999). *Advances in kernel methods – Support vector learning*. Cambridge, MA: MIT Press.

Tang, Z., & Fishwich, P. A. (1993). Back-propagation neural nets as models for time series forecasting. *ORSA Journal on Computing, 5*(4), 374–385.

Tay, F. E. H., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. *Omega: The International Journal of Management Science, 29*, 309–317.

Vapnik, V. N. (1998). *The nature of statistical learning theory*. New York: Springer-Verlag.

Wang, L. P. (Ed.). (2005). *Support Vector Machines: Theory and Application*. Berlin, Heidelberg, New York: Springer.

Wang, J. H., & Leu, J. Y. (1996). Stock market trend prediction using ARIMA-based neural networks. In *Proceedings of IEEE International Conference on Neural Networks 4*, 2160– 2165.

Yamada, H. (2005). Wavelet-based beta estimation and Japanese industrial stock prices. *Applied Economics Letters, 12*, 85–88.

Yao, J., & Tan, C. L. (2000). A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing, 34*, 79–98.

Zhang, G., & Hu, M. Y. (1998). Neural network forecasting of the British pound/US dollar exchange rate. *OMEGA: International Journal of Management Science, 26*, 495–506.

Zimmermann, H., Neuneier, R., & Grothmann, R. (2001). Multi-agent modeling of multiple FX-markets by neural networks. *IEEE Transactions on Neural Networks, 12*(4), 735–743.