Available online at www.sciencedirect.com
SCIENCE DIRECT®

Omega 33 (2005) 497 – 505

omega
The International Journal of Management Science

www.elsevier.com/locate/omega

A hybrid ARIMA and support vector machines model in stock price forecasting

Ping-Feng Pai*, Chih-Sheng Lin

Department of Industrial Engineering and Technology Management, Da-Yeh University, 112 Shan Jiau Road, Da-Tsuen, Chung-Hwa 515, Taiwan

Received 26 August 2003; accepted 22 July 2004
Available online 16 September 2004

Abstract

Traditionally, the autoregressive integrated moving average (ARIMA) model has been one of the most widely used linear models in time series forecasting. However, the ARIMA model cannot easily capture the nonlinear patterns. Support vector machines (SVMs), a novel neural network technique, have been successfully applied in solving nonlinear regression estimation problems. Therefore, this investigation proposes a hybrid methodology that exploits the unique strength of the ARIMA model and the SVMs model in forecasting stock prices problems. Real data sets of stock prices were used to examine the forecasting accuracy of the proposed model. The results of computational tests are very promising.

Keywords: Artificial neural networks; ARIMA; Support vector machines; Time series forecasting; Stock prices

1. Introduction

Forecasting stock prices has been regarded as one of the most challenging applications of modern time series forecasting. Thus, numerous models have been depicted to provide the investors with more precise predictions. Recently, artificial neural networks (ANN) have been applied to solve problems of forecasting stock prices. Kimoto and Asakawa [1] used modular neural networks to predict the timing of buying and selling for the Tokyo Stock Exchange. Experimental results showed that an excellent profit was achieved. Kamijo and Tanigawa [2] developed a pattern recognition technique to predict the stock prices on the Tokyo Stock Exchange. A new method has been presented to evaluate the recurrent networks to decrease the mismatching patterns. Yoon and Swales [3] presented a four-layered neural network to predict stock prices in the United States. The results revealed that the proposed approach outperforms the MDA (multiple discriminant analysis) method. Baba and Kozaki [4] presented a back-propagation neural network combined with a random optimization technique to predict stock markets in Japan. Simulation results proved that the proposed approach indeed helped to forecast stock prices. Cheung et al. [5] employed an adaptive rival penalized competitive learning method and a combined linear prediction model to forecast both the Shanghai share and the US Dollar to German Deutschmark exchange rate. Experimental results revealed that the proposed model increased profits with over time. Takahashi et al. [6] proposed a neural network that embodied a multiple line-segments regression technique to predict stock prices. The tangent and length of multiple line-segments regression were specified as outputs of the neural networks. The results showed that the proposed approach performed well in predicting stock prices. Kim and Chun [7] developed an arrayed probabilistic network with

* Corresponding author. Tel.: +886-4-8511888; fax: +886-4-8511270.
E-mail address: paipf@mail.dyu.edu.tw (P.-F. Pai).

0305-0483/$ - see front matter
doi:10.1016/j.omega.2004.07.024

a multiple value output model to predict a stock market index. Their approach outperforms the case based reasoning model, the recurrent neural network approach and the traditional back-propagation neural networks, in forecasting stock markets. Cristea and Okamoto [8] applied an approach that involved mathematical deductions of the energy function to Lyapunov Gradient Descent Neural Networks. Their proposed networks were used to predict stock markets. The results indicated that the approach outperforms traditional back-propagation neural networks and the random walk generator. Saad et al. [9] conducted a comparative investigation of TDNN (time-delay neural networks), PNN (probabilistic neural networks), and RNN (recurrent neural networks) in predicting daily closing prices in stock markets. The results showed that all networks are equally feasible but the most convenient network is preferable. Donaldson and Kamstra [10] applied multiplayer feedforward networks with nonlinear combinations to predict S&P the 500 stock index. The proposed model can account forthe effects of interactions between time series forecasts, and their approach outperforms other conventional forecasting approaches. Kim and Han [11] applied genetic algorithms to discriminate features in a backpropagation neural network. The proposed approach can determine connection weights of neural networks to predict the stock index. Simulation results indicated that the presented model outperforms linear transformation with backpropagation and linear transformation with the ANN trained by genetic algorithms. Oh and Kim [12] presented backpropagation neural networks that incorporated chaotic and piecewise techniques to deal with problems of predicting stock markets. The presented model is more profitable than the traditional backpropagation neural networks. Leigh et al. [13] combined pattern recognition with neural networks to predict the New York Stock Exchange Composite Index. The experimental results were encouraging and revealed the capabilities of the proposed hybrid model.

Different forecasting models can complement each other in capturing patterns of data sets, and both theatrical and empirical studies have concluded that a combination of forecast outperforms individual forecasting models [14–16]. Since the early work of Bates and Granger [17], several architectures of combined forecasts have been explored. Clemen [18] had a comprehensive bibliography review in this area. Menezes et al. [19] offered good guidelines for combined forecasting. They concluded that the problem of combined forecasts is implementing multi-criteria process and judging the attributes of an error specification. Lam et al. [20] proposed a goal programming model to obtain optimal weights for combining forecasting models. Terui and Dijk [21] presented a linear and nonlinear time series model for forecasting the US monthly employment rate and production indices. Their results demonstrated that the combined forecasts outperformed the individual forecasts. Fang [22] used quarterly UK consumption expenditure data to show the superiority of the combined forecasting model. Zhang [23] combined the ARIMA and feedforward neural networks models in forecasting. This study presents a hybrid model of ARIMA and SVMs to slove the stock price forecasting problem.

2. Hybrid model in forecasting

2.1. ARIMA model

Introduced by Box and Jenkins [24], the ARIMA model has been one of the most popular approaches to forecasting. In an ARIMA model, the future value of a variable is supposed to be a linear combination of past values and past errors, expressed as follows

$$\begin{aligned} y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} \\ + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}, \end{aligned} \tag{1}$$

where y_t is the actual value and ε_t is the random error at time t, ϕ_i and θ_j are the coefficients, p and q are integers that are often referred to as autoregressive and moving average polynomials, respectively. Basically, this method has three phases: model identification, parameter estimation and diagnostic checking. For example, the ARIMA(1,0,1) model can be represented as follows

$$y_t = \theta_0 + \phi_1 y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}. \tag{2}$$

The ARIMA model is basically a data-oriented approach that is adapted from the structure of the data themselves. However, any significant nonlinear data set limit the ARIMA. Therefore, the proposed hybrid model used the SVMs to deal with the nonlinear data pattern.

2.2. Support vector machines

The support vector machines (SVMs) were proposed by Vapnik [25]. 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

$$y = w\phi(x) + b, \tag{3}$$

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

$$R(C) = C\,\frac{1}{N}\sum_{i=1}^{N} L_\varepsilon(d_i, y_i) + \frac{1}{2}\,\|w\|^2, \tag{4}$$

$$L_\varepsilon(d, y) = \begin{cases} |d-y| - \varepsilon & |d-y| \geqslant \varepsilon, \\ 0 & \text{others}, \end{cases} \tag{5}$$

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 stock price in the ith period. This function indicates that errors below ε are not penalized. The term $C(1/N)\sum_{i=1}^{N} L_\varepsilon(d_i, y_i)$ is the empirical error. The second term, $\frac{1}{2}\|w\|^2$, measures the flatness of the function. C evaluates the trade-off between the empirical risk and the flatness of the model. Introducing the positive slack variables ζ and ζ^*, which represent the distance from the actual values to the corresponding boundary values of ε-tube. Eq. (4) is transformed to the following constrained formation:

Minimize :

$$R(w, \zeta, \zeta^*) = \frac{1}{2} ww^{\mathrm{T}} + C^* \left(\sum_{i=1}^{N} (\zeta_i + \zeta_i^*) \right) \tag{6}$$

Subjected to:

$$w\phi(x_i) + b_i - d_i \leqslant \varepsilon + \zeta_i^*, \tag{7}$$

$$d_i - w\phi(x_i) - b_i \leqslant \varepsilon + \zeta_i, \tag{8}$$

$$\zeta_i, \zeta_i^* \geqslant 0, \tag{9}$$

$$i = 1, 2, \ldots, N.$$

Finally, introducing Lagrangian multipliers and maximizing the dual function of Eq. (6) changes Eq. (6) to the following form:

$$\begin{aligned} R(\alpha_i - \alpha_i^*) = {} & \sum_{i=1}^{N} d_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) \\ & - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\alpha_i - \alpha_i^*) \\ & \times (\alpha_j - \alpha_j^*) K(x_i, x_j) \end{aligned} \tag{10}$$

with the constraints

$$\sum_{i=1}^{N} (\alpha_i - \alpha_i^*) = 0, \tag{11}$$

$$0 \leqslant \alpha_i \leqslant C, \tag{12}$$

$$0 \leqslant \alpha_i^* \leqslant C, \tag{13}$$

$$i = 1, 2, \ldots, N.$$

In Eq. (10), α_i and α_i^* are called Lagrangian multipliers. They satisfy the equalities,

$$\alpha_i * \alpha_i^* = 0,$$

$$f(x, \alpha, \alpha^*) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) K(x, x_i) + b. \tag{14}$$

Here, $K(x, x_i)$ is called the kernel function. The value of the kernel is equal to the inner product of two vectors x_i and x_j in the feature space $\phi(x_i)$ and $\phi(x_j)$, such that $K(x_i, x_j) = \phi(x_i) * \phi(x_j)$. Any function that satisfying Mercer's condition [25] can be used as the Kernel function. The Gaussian kernel function

$$K(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / (2\sigma^2))$$

is specified in this study. The SVMs were employed to estimate the nonlinear behavior of the forecasting data set because Gaussian kernels tend to give good performance under general smoothness assumptions.

2.3. The hybrid methodology

The behavior of stock prices can not easily be captured. Therefore, a hybrid strategy that has both linear and nonlinear modeling abilities is a good alternative for forecasting stock prices. Both the ARIMA and the SVMs models have different capabilities to capture data characteristics in linear or nonlinear domains, so the hybrid model proposed in this study is composed of the ARIMA component and the SVMs component. Thus, the hybrid model can model linear and nonlinear patterns with improved overall forecasting performance. The hybrid model (Z_t) can then be represented as follows

$$Z_t = Y_t + N_t, \tag{15}$$

where Y_t is the linear part and N_t is the nonlinear part of the hybrid model. Both Y_t and N_t are estimated from the data set. $\tilde{Y}_t$ is the forecast value of the ARIMA model at time t. Let ε_t represent the residual at time t as obtained from the ARIMA model; then

$$\varepsilon_t = Z_t - \tilde{Y}_t. \tag{16}$$

The residuals are modeled by the SVMs and can be represented as follows

$$\varepsilon_t = f(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-n}) + \Delta_t, \tag{17}$$

where f is a nonlinear function modeled by the SVMs and Δ_t is the random error. Therefore, the combined forecast is

$$\tilde{Z}_t = \tilde{Y}_t + \tilde{N}_t. \tag{18}$$

Notably, $\tilde{N}_t$ is the forecast value of (17).

3. Forecasting of stock prices

Ten stocks were used in this study to examine the performance of the proposed model. The daily closing prices of the stocks were collected. Fifty data (from Oct. 21, 2002 to Dec. 31, 2002) for each company were used as a training data set. Daily stock closing prices in January were used as a validation data set. Daily stock closing prices in February 2003 were used as a testing data set. In this study, only one-step-ahead forecasting is considered. One-step-ahead forecasting can prevent problems associated with cumulative errors from the previous period for out-of-estimation sample

Table 1
Data sets of stock prices

Training data set	Validation data set	Testing data set
10/21/2002 $\sim$ 1/31/2003	1/2/2003 $\sim$ 02/28/2003	02/03/2003 $\sim$ 12/31/2002

forecasting [26–30]. Table 1 lists the corresponding periods. Four indices, MAE (mean absolute error), MSE (mean square error), MAPE (mean absolute percent error), and RMSE (root mean square error), were used as measures of forecasting accuracy. The indices are shown as follows

$$\mathrm{MAE} = \frac{1}{N}\sum_{t=1}^{N} |d_i - z_t|, \tag{19}$$

$$\mathrm{MAPE} = \frac{100}{N}\sum_{t=1}^{N} \left|\frac{d_t - z_t}{d_t}\right|, \tag{20}$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{t=1}^{N} (d_t - z_t)^2, \tag{21}$$

$$\mathrm{RMSE} = \left\{\frac{1}{N}\sum_{t=1}^{N} (d_t - z_t)^2\right\}^{0.5}, \tag{22}$$

where N is the number of forecasting periods, d_i is the actual stock price at period t, and z_t is the forecasting stock price at period t.

In this study, the ARIMA model has three phases: model identification, parameter estimation, and diagnostic checking. Table 2 shows the most appropriate Box–Jenkins model for stock prices of different companies. ARIMA(0,1,0) models, which are random walk models, are suitable for predicting the stock prices of all the companies except for SBC Communications Inc. For the SVMs models, three parameters: σ, ε, and C, were adjusted based on the validation sets. The parameter sets with the lowest values of MSE were selected for use in the best fitted model. Of the hybrid models, ARIMA served as a preprocessor to filter the linear pattern of data sets. Then, the error terms from ARIMA were fed into the SVMs in the hybrid models. The SVMs were conducted to reduce the error function from the ARIMA. The three parameters (ε, C, and σ) of SVMs were adjusted. Improper selection of parameters can cause either over-fitting or under-fitting of the training data. Figs. 1–10 present MSE values of the hybrid model that correspond to the σ values. The same procedure was applied to the single SVMs models. Table 2 lists suitable parameters for different models.

Table 3 compares the forecasting results of different models. Those results indicate that the hybrid model outperforms the other two individual models (model 1 and 2) in terms of four indices, revealing that neither the ARIMA model

Table 2
Parameters of different models

Companies/ Models	(1) Eastman Kodak Company	(2) General Motors Corporation	(3) J.P. Morgan Chase & Co.	(4) Altria Group, Inc. (Philip Morris USA Inc.)	(5) SBC Communications Inc.	(6) Citigroup Inc.	(7) General Electric Company	(8) Southwest Water Company	(9) American National Insurance Company	(10) ATP Oil & Gas Corporation
ARIMA models	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	(1,0,0)	(0,1,0)	(0,1,0)	(0,1,0)	(0,1,0)	
SVMs models	$s = 0.3$	$s = 1.9$	$s = 1.3$	$s = 1.3$	$s = 1.5$	$s = 0.6$	$s = 0.6$	$s = 1.7$	$s = 1.4$	$s = 1.2$
	$= 0$	$= 0.3$	$= 0$	$= 0$	$= 0.1$	$= 0$	$= 0.4$	$= 0$	$= 0$	$= 0$
	$C = 100$	$C = 100$	$C = 100$	$C = 100$	$C = 100$	$C = 100$	$C = 100$	$C = 100$	$C = 100$	$C = 100$
Hybrid models	$s = 1$	$s = 3.4$	$s = 2$	$s = 4.1$	$s = 4.2$	$s = 1.0$	$s = 3.2$	$s = 0.3$	$s = 1.7$	$s = 2.0$
	$= 0.2$	$= 0$	$= 0$	$= 0$	$= 0$	$= 0.1$	$= 0.4$	$= 0.4$	$= 0.8$	$= 0.2$
	$C = 10$	$C = 1$	$C = 1$	$C = 1$	$C = 1$	$C = 10$	$C = 10$	$C = 10$	$C = 10$	$C = 10$

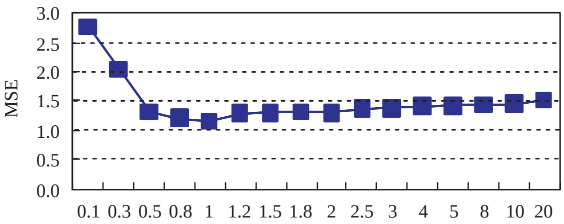

Fig. 1. MSE of Eastman Kodak Company ($\varepsilon = 0.2$, $C = 10$).

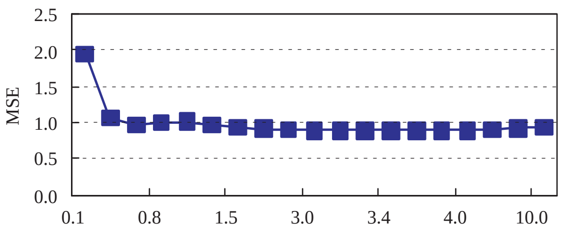

Fig. 2. MSE of General Motors Corporation ($\varepsilon = 0$, $C = 1$).

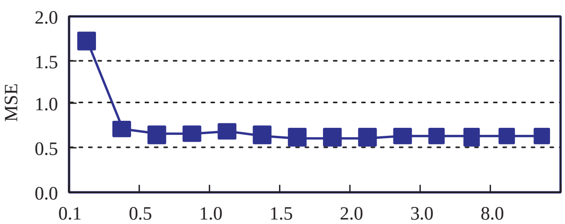

Fig. 3. MSE of J.P. Morgan Chase & Co. ($\varepsilon = 0$, $C = 1$).

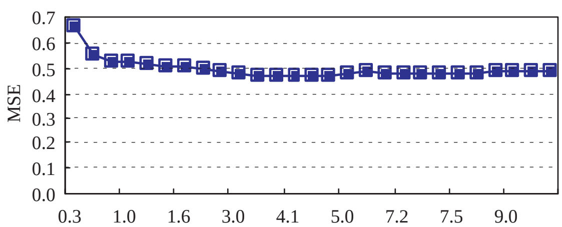

Fig. 4. MSE of Philip Morris Inc. ($\varepsilon = 0$, $C = 1$).

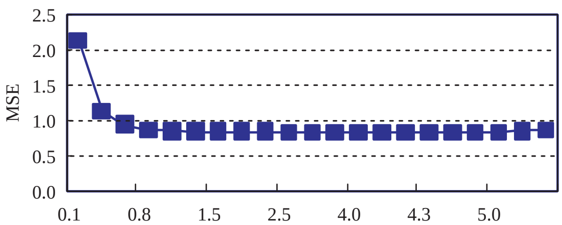

Fig. 5. MSE of SBC Communications Inc. ($\varepsilon = 0$, $C = 1$).

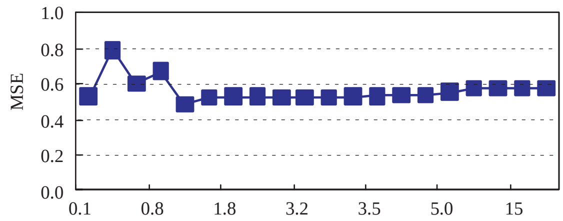

Fig. 6. MSE of Citigroup Inc. ($\varepsilon = 0.1$, $C = 10$).

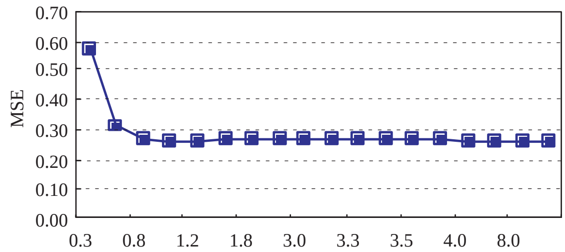

Fig. 7. MSE of General Electric Company ($\varepsilon = 0.4$, $C = 10$).

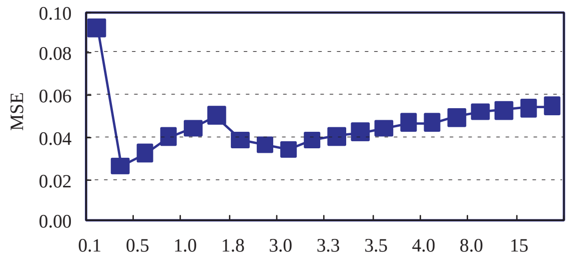

Fig. 8. MSE of Southwest Water Company ($\varepsilon = 0.4$, $C = 10$).

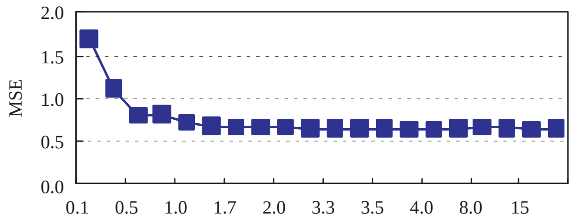

Fig. 9. MSE of American National Insurance Company ($\varepsilon = 0.8$, $C = 10$).

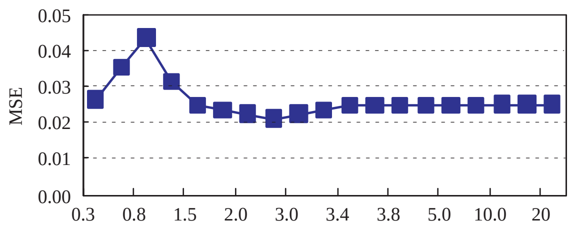

Fig. 10. MSE of ATP Oil & Gas Corporation ($\varepsilon = 0.2$, $C = 10$).

Table 3
Comparison of forecasting indices

	MAE	MSE	MAPE	RMSE
ARIMA model (*Model* 1)				
(1) Eastman Kodak Company	0.3495	0.2257	1.1494	0.4751
(2) General Motors Corporation	0.4905	0.3748	1.4214	0.6122
(3) J.P. Morgan Chase & Co.	0.2974	0.1284	1.3463	0.3584
(4) Philip Morris USA Inc.	0.3600	0.1873	0.9720	0.4328
(5) SBC Communications Inc.(ARIMA(1,0,0))	0.5628	0.4262	2.4843	0.6528
(5-1) SBC Communications Inc.(ARIMA(0,1,0))	0.5753	0.4499	2.5683	0.6708
(6) Citigroup Inc.	0.4879	0.3078	1.4892	0.5548
(7) General Electric Company	0.3079	0.1354	1.3214	0.3679
(8) Southwest Water Company	0.1189	0.02597	0.9127	0.1611
(9) American National Insurance Company	0.8079	1.1301	1.0051	1.0630
(10) ATP Oil & Gas Corporation	0.8211	0.0133	1.8915	0.1153
SVMs model (*Model* 2)				
(1) Eastman Kodak Company	0.3466	0.2247	1.1433	0.4740
(2) General Motors Corporation	0.4352	0.3186	1.2654	0.5644
(3) J.P. Morgan Chase & Co.	0.2980	0.1277	1.3501	0.3574
(4) Philip Morris USA Inc.	0.3603	0.1879	0.9731	0.4335
(5) SBC Communications Inc.	0.5445	0.4254	2.3954	0.6522
(6) Citigroup Inc.	0.5020	0.3337	1.5324	0.5777
(7) General Electric Company	0.3008	0.1336	1.2853	0.3655
(8) Southwest Water Company	0.1125	0.02537	0.8658	0.1593
(9) American National Insurance Company	0.8726	1.1126	1.0794	1.0548
(10) ATP Oil & Gas Corporation	0.7864	0.0119	1.8027	0.1089
Model 3: (*Model* 1 + *Model* 2)				
(1) Eastman Kodak Company	0.3499	0.2138	1.1534	0.4624
(2) General Motors Corporation	0.4586	0.3569	1.3391	0.5974
(3) J.P. Morgan Chase & Co.	0.2710	0.1095	1.2306	0.3310
(4) Philip Morris USA Inc.	0.3595	0.1924	0.9736	0.4386
(5) SBC Communications Inc.	0.6519	0.5890	2.9405	0.7674
(6) Citigroup Inc.	0.7175	0.7232	2.1928	0.8504
(7) General Electric Company	0.4230	0.2613	1.8229	0.5112
(8) Southwest Water Company	0.1224	0.02549	0.9330	0.1596
(9) American National Insurance Company	0.9120	1.4676	1.1359	1.2115
(10) ATP Oil & Gas Corporation	0.1276	0.0249	2.9447	0.1578
Hybrid model				
(1) Eastman Kodak Company	0.2303	0.1000	0.7598	0.3162
(2) General Motors Corporation	0.2579	0.2049	0.7550	0.4526
(3) J.P. Morgan Chase & Co.	0.2700	0.1125	1.2266	0.3354
(4) Philip Morris USA Inc.	0.2194	0.1101	0.5937	0.3319
(5) SBC Communications Inc.	0.1380	0.0742	0.6300	0.2725
(6) CitigroupInc.	0.4489	0.2626	1.3606	0.5125
(7) General Electric Company	0.2832	0.1324	1.2198	0.3639
(8) Southwest Water Company	0.1176	0.0251	0.9031	0.1584
(9) American National Insurance Company	0.7839	1.0027	0.9634	1.0014
(10) ATP Oil & Gas Corporation	0.0775	0.0114	1.7988	0.1069

nor the SVM model can capture all of the patterns in the data. The hybrid model is, however, can significantly reduce the overall forecasting errors. Furthermore, the model 3 is a combined model that uses the parameters in models 1 and 2. The results in Table 3 indicate the proposed hybrid model is superior to model 3. It is indicated that the combination of the best individual forecasting models does not necessarily yield favorable forecasting results. Figs. 11–20 make

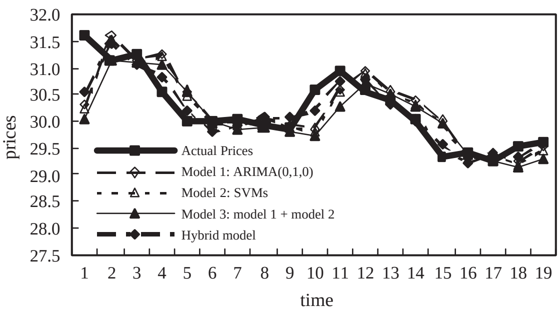

Fig. 11. Stock prices of Eastman Kodak Company.

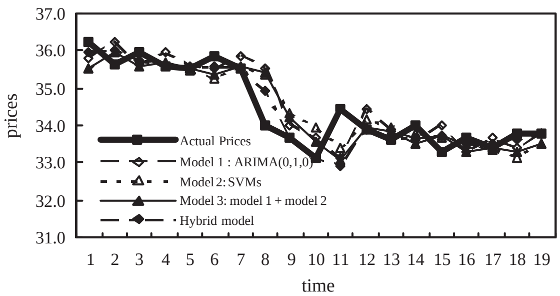

Fig. 12. Stock prices of General Motors Corporation.

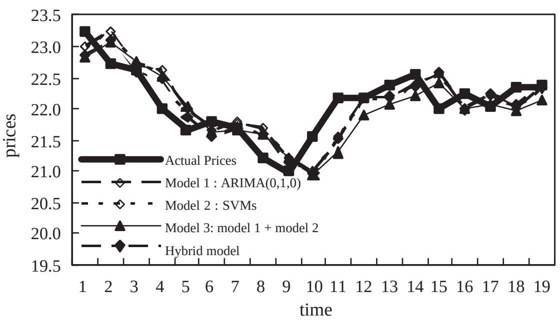

Fig. 13. Stock prices of J.P. Morgan Chase & Co.

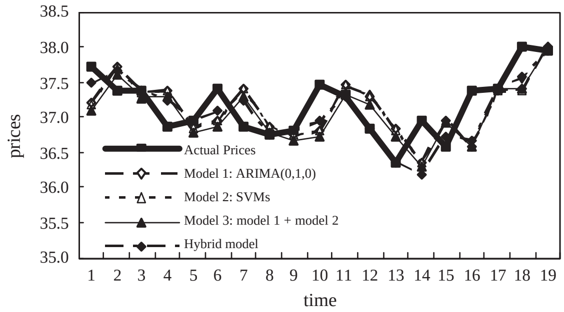

Fig. 14. Stock prices of Philip Morris Inc.

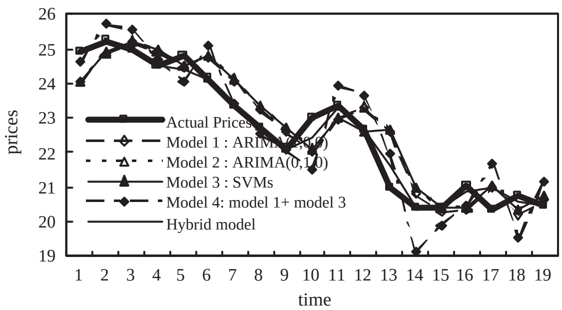

Fig. 15. Stock prices of SBC Communications Inc.

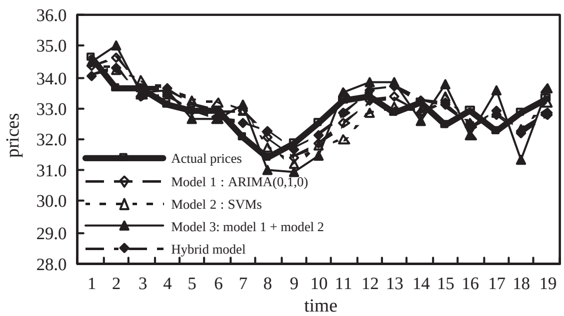

Fig. 16. Stock prices of Citigroup Inc.

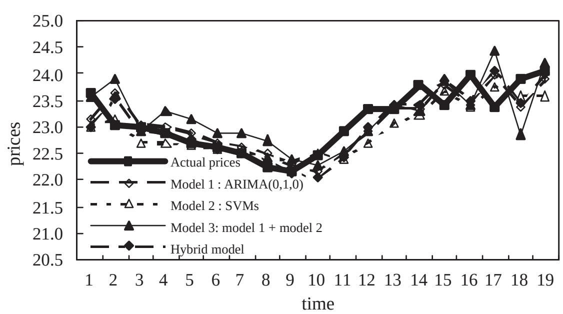

Fig. 17. Stock prices of General Electric Company.

point-to-point comparisons of actual and predicted values. Finally, an ARIMA (0,1,0) model was used to forecast the stock prices of the SBC Communications company to compare the forecasting benchmark with the random walk models . The results reveal that the proposed hybrid models yield better forecasting results than the random walk models.

4. Conclusions

For more than half a century, the autoregressive integrated moving average model has dominated many areas of time series forecasting. Recently, ANN has demonstrated the capability to capture the nonlinear data pattern. This study is motivated by evidence that different forecasting models can

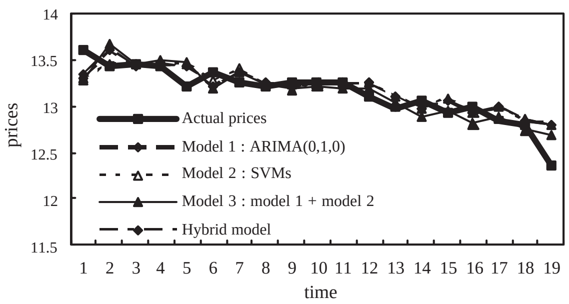

Fig. 18. Stock prices of Southwest Water Company.

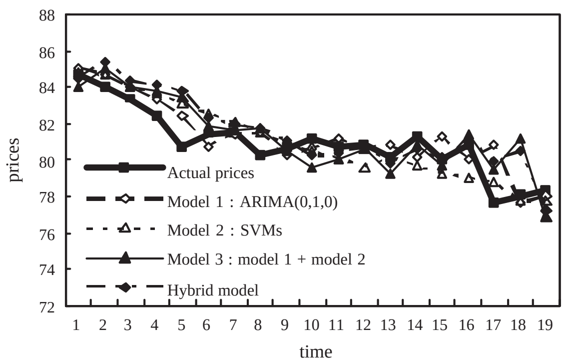

Fig. 19. Stock prices of American National Insurance Company.

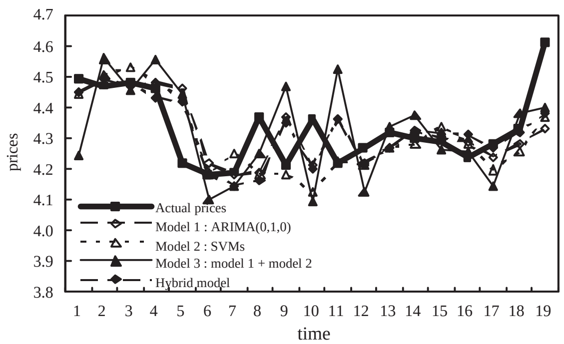

Fig. 20. Stock prices of ATP Oil & Gas Corporation.

complement each other in approximating data sets, and proposed a hybrid model of the ARIMA and the SVMs. The presented model is believed to greatly improve the prediction performance of the single ARIMA model or the single SVMs model in forecasting stock prices. Theoretically as well as empirically, hybridizing two dissimilar models reduces forecasting errors [31,32]. However, future research should address some problems. This study demonstrated that a simple combination of the two best individual models does not necessarily produce the best results. Therefore, the structured selection of optimal parameters of the hybrid model is of great interest.

Acknowledgements

This research was conducted with the support of National Science Council, Taiwan, R.O.C. (NSC92-2213-E-212-001 & NSC 93-2213-E-212-001).

References

[1] Kimoto T, Asakawa K. Stock market prediction system with modular neural networks. In: Proceedings of the international joint conference on neural networks. 1990. p. 1–6.

[2] Kamijo K, Tanigawa T. Stock price pattern recognition—a recurrent neural network approach. In: Proceedings of the international joint conference on neural networks. 1990. p. 211–5.

[3] Yoon Y, Swalves G. Predicting stock price performance: a neural network approach. In: Proceedings of the twenty-fourth annual Hawaii international conference on system sciences. 1991. p. 156–62.

[4] Baba N, Kozaki M. An intelligent forecasting system of stock price using neural networks. In: Proceedings of the international joint conference on neural networks. 1992. p. 371–7.

[5] Cheung YM, Lai HZH, Xu L. Application of adaptive RPCL-CLP with trading system to foreign exchange investment. In: Proceedings of the IEEE international conference on neural networks. 1996. p. 2033–8.

[6] Takahashi T, Tamada R, Nagasaka K. Multiple line-segments regression for stock prices and long-range forecasting system by neural networks. In: Proceedings of the 37th SICE annual conference. 1998. p. 1127–32.

[7] Kim SH, Chun SH. Graded forecasting using an array of bipolar predictions: application of probabilistic neural to a stock market index. International Journal of Forecasting 1998;14:323–37.

[8] Cristea AI, Okamoto T. Energy function construction and implementation for stock exchange prediction NNs. In: Proceedings of the international conference on knowledge-based intelligent electronic systems. 1998. p. 403–10.

[9] Saad EW, Prokhorov DV, Wunsch DC. Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks. IEEE Transactions on Neural Networks 1998;9:1456–70.

[10] Donaldson RG, Kamstra M. Neural network forecast combining with interaction effects. Journal of the Franklin Institute 1999;336:227–36.

[11] Kim KJ, Han I. Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. Expert Systems with Application 2000;19:125–32.

[12] Oh KJ, Kim KJ. Analyzing stock market tick data using piecewise nonlinear model. Expert Systems with Application 2002;22:249–55.

[13] Leigh W, Paz M, Purvis R. An analysis of a hybrid neural network and pattern recognition technique for predicting short-term increases in the NYSE composite index. The International Journal of Management Science 2002;30:69–76.

[14] Makridakis S, Winkler RL. Averages of forecast. Management Science 1983;29:987–96.

[15] Lawerence MJ, Edmundson RH, O'Connor MJ. The accuracy of combining judgemental and stastical forecasts. Management Science 1986;32:1521–32.
[16] Makridakis S. Why combining works?. International Journal of Forecasting 1989;5:601–3.
[17] Bates JM, Granger CWJ. Combination of forecasts. Operational Research Quarterly 1969;20:451–68.
[18] Clemen R. Combining forecasts: a review and annotated bibliography with discussion. International Journal of Forecasting 1989;5:559–608.
[19] Menezes Lilian M, Bunn Derek W, Taylor James W. Review of guidelines for the use of combined forecasts. European Journal of Operational Research 2000;120:190–204.
[20] Lam KF, Mui HW, Yuen HK. A note on minimizing absolute percentage error in combined forecasts. Computer & Operations Research 2001;28:1141–7.
[21] Nobuhiko Terui, Dijk Herman K. Combined forecasts from linear and nonlinear time series models. International Journal of Forecasting 2002;18:421–38.
[22] Fang Y. Forecasting combination and encompassing tests. International Journal of Forecasting 2003;19:87–94.
[23] Zhang G. Peter Times series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 2003;50:159–75.
[24] Box GEP, Jenkins G. Time series analysis, forecasting and control. San Francisco, CA: Holden-Day; 1970.
[25] Vapnik V. The nature of statistic learning theory. New York: Springer; 1995.
[26] Makridakis S, Winkler RL. Sampling distribution of post-sample forecasting errors. Applied Statistics 1989;38:331–42.
[27] Pack DJ. In defense of ARIMA modeling. International Journal of Forecasting 1990;6:211–8.
[28] Vald_erio AR, Silvia L. Some simulations and applications of forecasting long memory time series models. Journal of Statistical Planning and Inference 1999;80:269–87.
[29] Sivakumar B, Jayawardena AW, Fernando TMKG. River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches. Journal of Hydrology 2002;265:225–45.
[30] McAvinchey ID, Yannopoulos A. Stationary, structural change and specification in a demand system: the case of energy. Energy Economics 2003;25:65–92.
[31] Granger CWJ. Combining forecasts—twenty years later. Journal of Forecasting 1989;8:167–73.
[32] Krogh A, Vedelsby J. Neural network ensembles, cross validation, and active learning. Advanced Neural Information Process 1995; 231–8.