Paper

Application of Support Vector Machine to Forex Monitoring

Joarder Kamruzzaman* Non-member
Ruhul A Sarker** Non-member

Previous studies have demonstrated superior performance of artificial neural network (ANN) based forex forecasting models over traditional regression models. This paper applies support vector machines to build a forecasting model from the historical data using six simple technical indicators and presents a comparison with an ANN based model trained by scaled conjugate gradient (SCG) learning algorithm. The models are evaluated and compared on the basis of five commonly used performance metrics that measure closeness of prediction as well as correctness in directional change. Forecasting results of six different currencies against Australian dollar reveal superior performance of SVM model using simple linear kernel over ANN-SCG model in terms of all the evaluation metrics. The effect of SVM parameter selection on prediction performance is also investigated and analyzed.

Keywords: Forecasting, foreign exchange, neural network, support vector machine.

1. Introduction

The foreign exchange rates play an important role in the international trade and business. As a result, the appropriate prediction of exchange rate is a crucial factor for the success of many businesses and fund managers. It is generally known that the currency exchange market is highly volatile and fluctuating with time. Therefore, it is not possible to exploit or predict its behaviour. A common myth of traders, however, is that there is certain predictability in the currency market dynamics. There are several recent works in the field of prediction and forecasting which indicate that these beliefs by traders are real and the market in some aspect shows predictable dynamics [(2)(16)~(18)].

Prediction of exchange rates is one of the challenging applications of modern time series forecasting. The rates are inherently noisy, non-stationary and deterministically chaotic [(6)(33)]. These characteristics suggest that there is no complete information that could be obtained from the past behaviour of such markets to fully capture the dependency between the future rates and that of the past. One general assumption is made in such situations is that the historical data incorporate all those behaviour. That means the historical data is the major player in the prediction process. Under this condition, the obvious question is: how good are those predictions?

For the last two decades, Box and Jenkins' Auto-Regressive Integrated Moving Average (ARIMA) technique [(1)] has been widely used for time series forecasting. Because of its popularity, the ARIMA model has been used as a benchmark to evaluate some new modelling approaches [(14)]. However, ARIMA is a general univariate model which is developed based on the assumption that the time series being forecasted are linear and stationary [(2)].

Artificial Neural Network (ANN), the well-known function approximator for prediction and system modelling, has proved itself as a great tool for many applicability in time-series analysis and forecasting [(31)~(34)]. ANNs are universal function approximators that can map any nonlinear function without a priori assumptions about the data [(2)]. In several applications, Tang and Fishwich [(27)], Jhee and Lee [(15)], Wang and Leu [(29)], Hill *et al.* [(13)], and Shazly *et al.* [(26)], Leung *et al.* [(22)], and many other researchers have shown that ANNs perform better than ARIMA models, specifically, for more irregular series and for multiple-period-ahead forecasting. Kaastra and Boyd [(16)] provided a general introduction of how a neural network model should be developed to model financial and economic time series. Zhang and Hu [(34)] analysed backpropagation neural networks' ability to forecast an exchange rate. Wang [(30)] cautioned against the dangers of one-shot analysis since the inherent nature of data could vary. Klein and Rossin [(19)] proved that the quality of the data also affects the predictive accuracy of a model. More recently, Yao *et al.* [(31)] evaluated the capability of a backpropagation neural network model as an option price forecasting tool. They have recognised the fact that neural network models are context sensitive and when studies of this type are conducted, it should be as comprehensive as possible for different markets and different neural network models. Kamruzzaman and Sarker [(17)(18)] investigated three ANN based forecasting models using Standard Backpropagation, Scaled Conjugate Gradient and Backpropagation with Baysian Regularization with six technical indicators to predict six major currencies against Australian dollar. Performance measured in terms of five metrics showed that the best result is achieved with

* Faculty of Information Technology, Monash University, Gippsland Campus, Victoria 3842, Australia.
** School of IT & EE, University of NSW, ADFA Campus, NorthCott Drive, Canberra 2600, Australia.

scaled conjugate gradient model and all ANN based models outperform the ARIMA model.

Recent research has been directed to alternative techniques, particularly based on statistical learning theory. Support Vector Machine (SVM) has emerged as a new and powerful technique for learning from data and in particular for solving classification and regression problems with better performance [2][8]~[10][28]. The main advantage of SVM is its ability to minimize structural risk as opposed to empirical risk minimization as employed by the neural network system. Cao *et al.* [2] used SVM to forecast S & P 500 daily index in the Chicago Mercantile. In this study, we used SVM model to forecast foreign currency exchange rates of six different currencies against Australian dollar. Our work differs substantially from Cao's work in several aspects. One of the aims of this study was to investigate whether SVM can predict exchange rates with acceptably high accuracy because many authors claim "exchange rates are rather unpredictable, and that a random walk model is often a better predictor than concurrent nonlinear models" [2][3][7][24][25]. Cao *et al.* [2] used nine rather complex technical indicators while we were interested in achieving high accuracy using more simple technical indicators. In their work, performance of an SVM model was compared with that of an ANN model trained by standard Backpropagation (BP) algorithm. But it has been demonstrated in numerous studies that standard BP is extremely slow and does not yield good generalization ability in many cases. The present study compares SVM model's performance with an ANN model trained with scaled conjugate gradient algorithm which was found to perform best among the NN models studied earlier [17]. The proposed model was used to predict currency exchange rates of Australian Dollar with six other currencies such as US Dollar (USD), Great Britain Pound (GBP), Japanese Yen (JPY), Singapore Dollar (SGD), New Zealand Dollar (NZD) and Swiss Franc (CHF). A total 500 weeks (closing rate of the week) data were used to build the model and 65 weeks data to evaluate the models. The models were compared on the basis of five performance metrics widely used in time series forecasting.

Section 2 describes the theory of SVM regression, and Section 3 presents forecasting model and performance metrics. Section 4 describes the experimental results and the comparison with NN based model and section 5 presents conclusion.

2. Support Vector Machine Regression Model

Support vector machine introduced by Vapnik [28] has attracted much research attention in recent years due its demonstrated improved generalization performance over other techniques in many real world applications. It has been used in classification as well as regression tasks. The main difference between this technique and many other conventional regression techniques including neural network is that it minimizes the structural risk instead of the empirical risk. The principal is based of the fact that minimizing an upper bound on the generalization error rather than minimizing the training error is expected to perform better. The generalization error rate is bounded by the sum of training error rate and a term that depends on Vapnik-Chervonenkis (VC) dimension [12]. VC dimension is a measure of complexity of the dimension space. Support vector machines find a balance between the empirical error and the VC-confidence interval. SVMs perform by nonlinearly mapping the input data into a high dimensional feature space by means of a kernel function and then do the linear regression in the transformed space. The whole process results in nonlinear regression in the low-dimensional space.

Consider a data set consisting $G = (\mathbf{x}_i, y_i)_{i=1}^{N}$ of N data points where each input $\mathbf{x}_i$ is mapped into the corresponding output y_i. Given that the data set realizes some unknown function $g(\mathbf{x})$, we need to determine a function f that approximates $g(\mathbf{x})$, based on the knowledge of data set G. In SVM, $\mathbf{x}_i$ is first mapped into a higher dimensional space F via a nonlinear mapping, and perform linear regression in this space. The SVM approximates the function as

$$f(\mathbf{x}) = \sum_{i=1}^{D} \omega_i \Phi_i(\mathbf{x}) + b \quad \text{with} \quad \Phi : \Re^n \to F, \omega \in F, \qquad \cdots\cdots (1)$$

where ω_i are the coefficients and b is a threshold value. This approximation can be considered as a hyperplane in the D-dimensional feature space F defined by the functions $\Phi_i(\mathbf{x})$ where the dimensionality can be very high, possibly infinite. Since Φ is fixed, ω_i can be determined from the data by minimizing the sum of empirical risk and a complexity term defined in the following risk function.

$$R = C\frac{1}{N}\sum_{i=1}^{N} |y_i - f(\mathbf{x}_i)|_\epsilon + \frac{1}{2} \parallel \omega \parallel^2, \cdots\cdots (2)$$

where ϵ is a parameter to be set *a priori* and an error below ϵ is not penalized according to the following error function.

$$|y_i - f(\mathbf{x}_i)|_\epsilon = \begin{cases} 0 & \text{if} |y_i - f(\mathbf{x}_i)| < \epsilon \\ |y_i - f(\mathbf{x}_i)| & \text{otherwise.} \end{cases} \qquad \cdots\cdots (3)$$

The first term in Eq. (2) describes the ϵ-insensitive loss function and the second term is a measure of function flatness. Thus SVMs perform linear regression in high dimensional feature space using ϵ-insensitive loss and at the same time tries to reduce model complexity by minimizing $\|\omega\|^2$. The constant $C > 0$ is a regularization constant determining the trade-off between the training error and model flatness. Introducing the slack variables ξ and ξ^*, SVM regression is formulated as minimization of the following optimization problem:

$$\text{minimize} \quad \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \cdots\cdots (4)$$

$$\text{subject to} \begin{cases} f(\mathbf{x}_i) + b - y_i \leq \epsilon + \xi_i \\ y_i - f(\mathbf{x}_i) - b \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

The above optimization with constraint is a quadratic programming problem that can be solved by constructing a Lagrangian and transforming into the dual problem, and its solution is given as follows [28].

$$f(\mathbf{x}, \alpha, \alpha^*) = \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) K(\mathbf{x}, \mathbf{x}_i) + b, \cdots\cdots (5)$$

where Lagrange multipliers α_i and α_i^* are associated with each data point $\mathbf{x}_i$, and subject to the constraints $0 \leq \alpha_i^*, \alpha_i \leq C$ and $\sum_{i=1}^{N}(\alpha_i^* - \alpha_i) = 0$. Training points with nonzero Lagrange multipliers are called support vectors. The smaller the fraction of support vectors, the more general the solution is but large support vectors do not necessarily mean an overtrained solution. The kernel function *K(.)* describes an inner product in the *D*-dimensional space as below and satisfies the Mercer's condition [5].

$$K(\mathbf{x}, \mathbf{x}_i) = \sum_{i=1}^{D} \Phi_i(\mathbf{x}) \Phi_i(\mathbf{x}_i) \cdots\cdots\cdots\cdots\cdots\cdots (6)$$

The coefficients α, α^* are obtained by maximizing the following quadratic form subject to the conditions stated earlier:

$$R(\alpha, \alpha^*) = \sum_{i=1}^{N} y_i(\alpha_i^* - \alpha_i) - \epsilon \sum_{i=1}^{N} (\alpha_i^* + \alpha_i) - \frac{1}{2} \sum_{i,j=1}^{N} (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(\mathbf{x}_i, \mathbf{x}_j) \cdots\cdots\cdots\cdots\cdots\cdots (7)$$

Once the coefficients are determined, the regression estimate is given by Eq. (5). The threshold b is computed from the constraints in Eq. (4) using the fact that the first constraint becomes an equality with $\xi_i = 0$ if $0 < \alpha_i < C$, and the second constraint becomes an equality with ξ_i^* if $0 < \alpha_i^* < C$.

The generalization performance (i.e., accuracy in predicting exchange rates in this study) depends on the parameters C, ϵ and kernel type. Common kernel types are linear, polynomial and Gaussian kernels.

3. Forecasting Model and Evaluation

3.1 Forecasting Model Two major methodologies for financial forecasting are technical and fundamental analyses. Technical analysis has drawn particular academic interest due to the increasing evidence that markets are less efficient than was originally thought [21]. Like many other economic time series model, exchange rate exhibits its own trend, cycle, season and irregularity. As Medeiros *et al.* [24] argues, there are evidences that favour linear and nonlinear models against random walk, and nonlinear models stand a better chance when nonlinearity is spread in time series. This suggests that an SVM based model could predict the exchange rate accurately. In this study, we used time delay moving average as technical data. The advantage of moving average is its tendency to smooth out some of the irregularity that exits between market days [32]. In this model, previous week's closing exchange rate and five moving average values over the past 24 weeks, six technical indicators in total are used to build SVM model to predict following week's rate. The indicators are MA5 (moving average on 5 days), MA10, MA20, MA60, MA120 and X_i, namely, moving average of one week, two weeks, one month, one quarter, half year and last week's closing rate, respectively. The predicted value X_{i+1} is the following week's exchange rate.

3.2 Evaluation of the Model The forecasting performance of the above model is evaluated against five widely used statistical metrics, namely, Normalized Mean Square Error (NMSE), Mean Absolute Error (MAE), Directional Symmetry (DS), Correct Up trend (CP) and Correct Down trend (CD). These criteria are defined in Table 1. NMSE and MAE measure the deviation between the actual and forecasted values. Smaller values of these metrics indicate higher accuracy in forecasting. DS measures correctness in predicted directions. CP and CD measure the correctness of predicted up and down trend, respectively. In general, it is desirable for a forecasting model to attain low NMSE and MAE but high DS, CP and CD. In practice, we may have a model that yields superior NMSE and MAE but inferior DS, CP and CD than other models, or *vice versa.* Several authors have argued that directional change metrics may be a better standard for determining the quality of forecasting. However, which metric should be given more importance depends on trading strategies and how the prediction is used for trading advantages [31].

4. Simulation Results and Discussion

The foreign exchange rates of six different currencies against Australian dollar are used to build the model. The data showing the exchange rate of US dollar, British Pound, Japanese Yen, Singapore dollar, New Zealand dollar and Swiss Franc against Australian dollar during January 1991 to July 2000 is made available by the Reserve Bank of Australia. In total 565 weekly data were considered of which the first 500 weekly data were used as the training set to build the SVM model. The remaining 65 weekly data were used to validate the model.

To build an SVM model the user needs to select the kernel type and the value of the parameters C and ϵ. In this study, the model was build using the simple linear kernel defined as $K(\mathbf{x}_i, \mathbf{x}_j) = <\mathbf{x}_i \cdot \mathbf{x}_j>$ where $< \cdot >$ is the inner product operator. The regularization parameter C was varied from 0.1 to 10^5 and ϵ parameter was varied from 0.0005 to 0.006 to investigate their effects on prediction and discussed later in this section. The forecasted result over the test data was compared with the Scaled Conjugate Gradient (SCG) neural network model studied by Kamruzzaman and Sarker [17][18]. In that study, it was concluded that the scaled conjugate gradient based model was the best among the three neural network models tested for predicting those currency exchange rates under consideration. The performance metrics were calculated over 35 weeks and 65 weeks forecasted data as well as over the training data set. Table 2 shows the performance metrics over the test and training data set for six currencies against Australian dollar and presents a comparison with the neural net model. The table presents the best result achieved in each case by SVM model with C varied within the range specified above and keeping ϵ fixed at 0.001.

A comparison of results with neural network based model illustrates that SVM based forecasting model can predict the exchange rate better, in terms of both accuracy (NMSE, MAE) and trend prediction (DS, CP, CD). SVM based model produces lower NMSE and MAE and higher DS, CP and CD than those of NN based model. Especially in the first 35 weeks, prediction by SVM is much better than those achieved by the neural network

Table 1. Performance metrics used in this experiment

NMSE	$NMSE = \frac{\sum_k (x_k - \hat{x}_k)^2}{\sum_k (x_k - \bar{x}_k)^2} = \frac{1}{\sigma^2 N} \sum_k (x_k - \hat{x}_k)^2$
MAE	$MAE = \frac{1}{N} \lvert x_k - \hat{x}_k \rvert$
DS	$DS = \frac{100}{N} \sum_k d_k,$ $d_k = \begin{cases} 1 & \text{if } (x_k - x_{k-1})(\hat{x}_k - \hat{x}_{k-1}) \geq 0 \\ 0 & \text{otherwise.} \end{cases}$
CP	$CP = 100 \frac{\sum_k d_k}{\sum_k t_k},$ $d_k = \begin{cases} 1 & \text{if } (\hat{x}_k - \hat{x}_{k-1}) > 0, (x_k - x_{k-1})(\hat{x}_k - \hat{x}_{k-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases},$ $t_k = \begin{cases} 1 & \text{if } (x_k - x_{k-1}) > 0 \\ 0 & \text{otherwise.} \end{cases}$
CD	$CD = 100 \frac{\sum_k d_k}{\sum_k t_k},$ $d_k = \begin{cases} 1 & \text{if } (\hat{x}_k - \hat{x}_{k-1}) < 0, (x_k - x_{k-1})(\hat{x}_k - \hat{x}_{k-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases},$ $t_k = \begin{cases} 1 & \text{if } (x_k - x_{k-1}) < 0 \\ 0 & \text{otherwise.} \end{cases}$

Table 2. Performance metrics calculated over the training and test data for six currencies against Australian dollar

Period	Criteria	GB Pound		US Dollar		Japanese Yen		Singapore Dollar		NZ Dollar		Swiss Franc	
		NN	SVM	NN	SVM	NN	SVM	NN	SVM	NN	SVM	NN	SVM
35 week forecast	NMSE	0.1578	0.1276	0.2624	0.2198	0.1264	0.108	0.2321	0.1512	0.0878	0.0880	0.0485	0.0442
	MAE	0.0030	0.0027	0.0035	0.0031	0.6243	0.569	0.0076	0.0068	0.0038	0.0037	0.0059	0.0054
	DS	77.143	80.00	80.00	80.00	80.00	80.00	82.857	85.714	85.714	82.857	82.857	80.00
	CP	81.25	87.50	82.353	82.353	81.818	81.818	82.353	88.235	87.500	87.50	80.00	75.00
	CD	73.684	73.684	82.353	82.353	76.923	76.923	83.333	83.333	84.210	78.947	86.666	86.667
65 week forecast	NMSE	0.0729	0.0573	0.0418	0.0380	0.0411	0.0350	0.0760	0.0507	0.0217	0.0210	0.0389	0.0356
	MAE	0.0023	0.0020	0.0029	0.0028	0.5188	0.4789	0.0060	0.0053	0.0033	0.0031	0.0052	0.0048
	DS	84.615	84.615	81.538	81.538	81.538	81.538	86.154	87.692	84.615	83.077	84.615	83.077
	CP	87.879	87.879	78.947	78.947	83.784	83.784	88.235	91.176	82.143	82.143	82.857	80.000
	CD	83.871	83.871	88.461	88.461	78.571	78.571	83.871	83.871	88.889	86.111	86.666	86.667
Training Data	NMSE	0.0052	0.0052	0.0023	0.0024	0.0035	0.0033	0.0051	0.0052	0.0034	0.0037	0.2753	0.2860
	MAE	0.0025	0.0025	0.0027	0.0027	0.5454	0.5145	0.0047	0.0047	0.0039	0.0040	0.0113	0.0090
	DS	76.800	76.800	78.800	79.200	75.200	74.600	76.600	77.200	77.400	78.400	78.80	78.20
	CP	76.623	76.190	77.381	77.381	77.291	76.892	75.424	75.423	77.049	77.459	78.632	79.487
	CD	77.528	77.903	80.894	81.707	73.684	72.874	79.150	80.309	78.656	80.237	79.245	77.359

model. This suggests that further improvement in forecasting may be obtained by rebuilding the model probably after every 30 weeks. Better prediction performance by SVM was consistently observed in all currencies. For example, in forecasting Singapore dollar, NMSE produced by neural network over 35 weeks is 0.2321 and that by SVM is 0.1512 which is large improvement over neural network. Significant improvement by SVM is also shown in case of US dollar and British Pound, however, improvement is little in New Zealand dollar. This suggests that the amount of improvement depends on other factors including the pattern of the data. An analysis of the historical data of all the currencies shows that Japanese Yen experienced the most fluctuation within the period we considered in the training phase. On the other hand New Zealand dollar is the least fluctuating one. This suggests that, with much fluctuation in data, neural network is unable to find appropriate set of weight that can accurately forecast the exchange rate. However, with low fluctuation, it is easier for neural network to learn the training data. This makes SVM is more effective for predicting exchange rate when the training data is much fluctuating. Figure 1(a)–(f) shows the actual and forecasted time series by VSM and scaled conjugate gradient based neural network. For example, in predicting exchange rate of British Pound between weeks 18∼24 where a sudden sharp fall is followed by a steep rise (Fig.1(a)), SVM model could closely follow the actual rate whereas NN model suffered from gross deviation. Similar trend is also evident in other currencies.

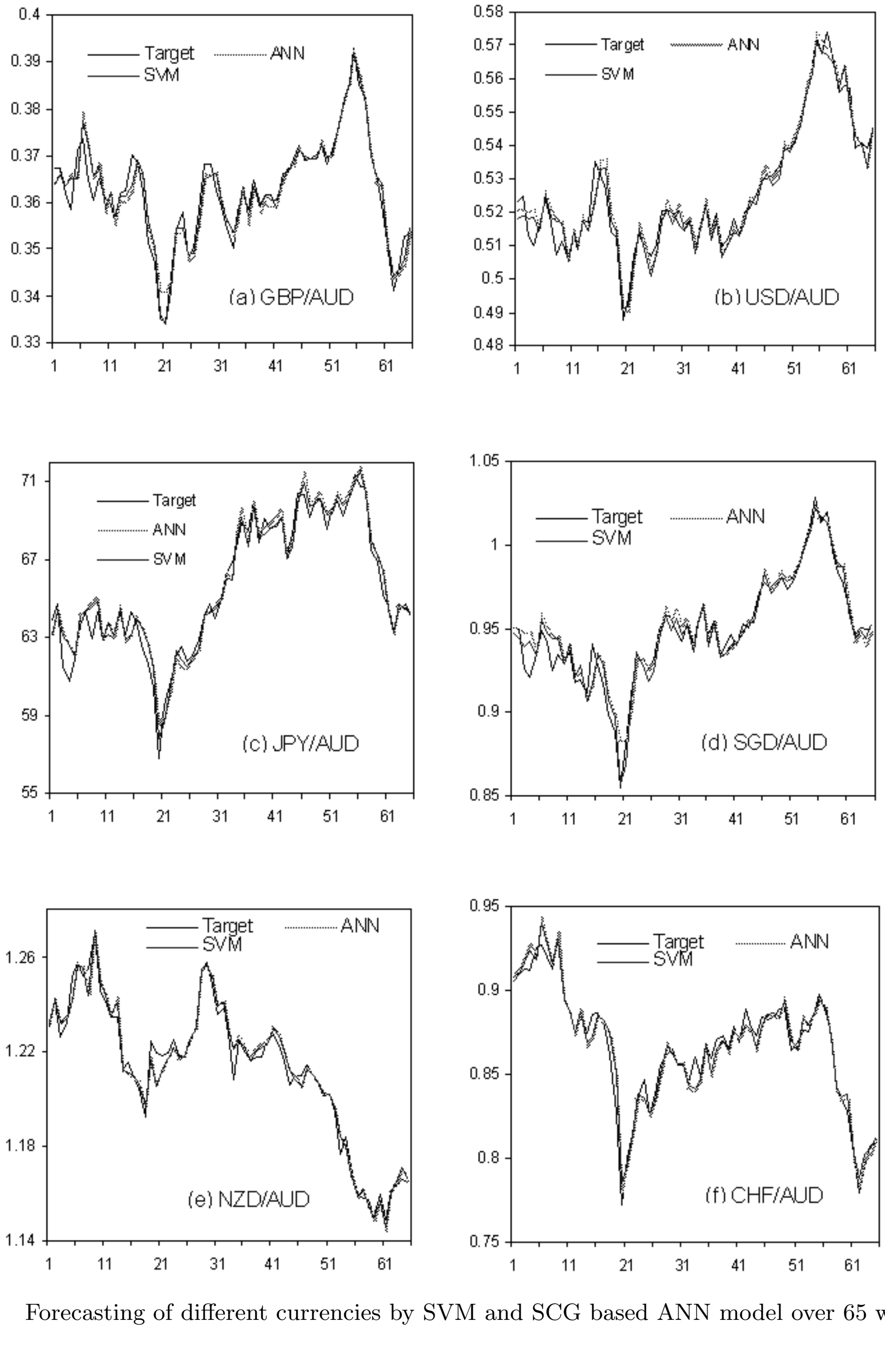

Fig. 1. Forecasting of different currencies by SVM and SCG based ANN model over 65 weeks.

The ability of SVM model to closely follow the actual rate in such a situation will make a significant difference in the outcome of financial gain in actual exchange trading.

In order to investigate the effect of regularization parameter C on prediction performance, we have experimented with different values of C ranging from 0.1 to 10^5. In general when C is too large, emphasis is placed on minimizing the empirical risk which may result in over-fitting of training data and lead to a deterioration of prediction performance. Too small value of C means the model will not learn from the training data properly resulting in under-fitting and leading to inferior performance. Figure 2(a)–(c) illustrates the prediction error of Japanese Yen versus C. Figure 2(a) shows that NMSE over 65 weeks forecasting result drops from 0.06 to 0.035 when C is increased from 0.5 to 50. For any value larger than 50, NMSE remains the same. MAE also exhibits the similar behaviour. Figure 2(c) shows that DS varies from 75.4% to 83.0% when C is increased from 0.5 to 7.0, later DS remains 81.54% as C increases further. This indicates little impact on the variation of prediction error with respect to regularization parameter at this stage. This kind of behaviour, i.e., negligible effect of C on performance when C is larger than a particular threshold is also supported in a recent study by Cherkassky *et al.* [(4)].

The performance of SVM also depends on the parameter ϵ which controls the width of ϵ-insensitive tube used to fit the training data. A larger value means support vectors will allow more margin of error while a smaller value will tightly fit the training data set but generalize poorly on the unseen data. It can also affect the number of support vectors that constructs the regression function. Larger ϵ-value results in fewer SVs and more ‘flat’ (less complex) regression estimates [(4)]. In order to investigate the effect of ϵ on prediction performance in this study, we experimented with 5 different values within the range 0.0005 to 0.006. Usual practice is to keep the value of ϵ small, however, the most appropriate value would depend on the particular data set. Figure 3 shows the effect of ϵ on NSME, MAE and DS. It displays that the best result is achieved when ϵ=0.001. All the three metrics deteriorate for any smaller or larger value of ϵ. However, even for relatively large value (0.006) SVM

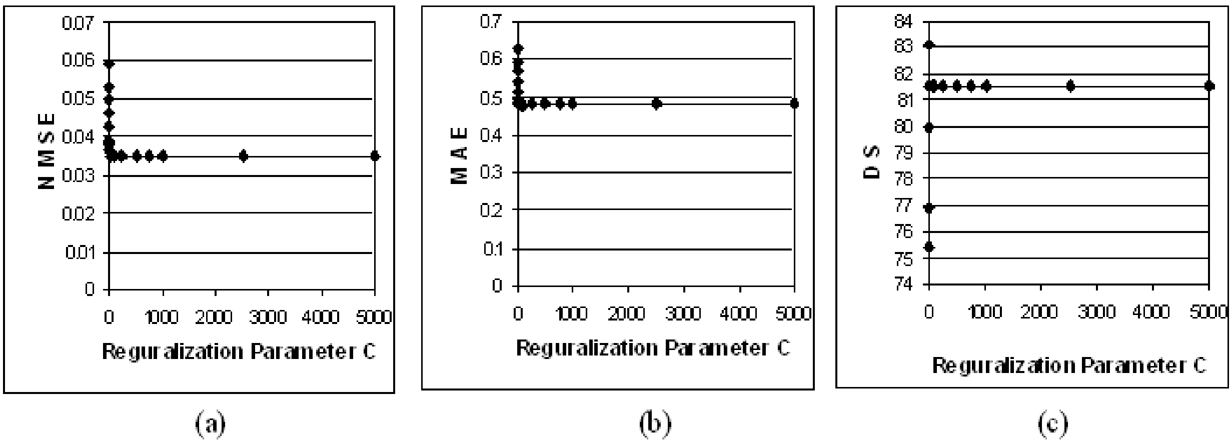

Fig. 2. Prediction performance versus regularization parameter C for predicting Japanese Yen exchange rate.

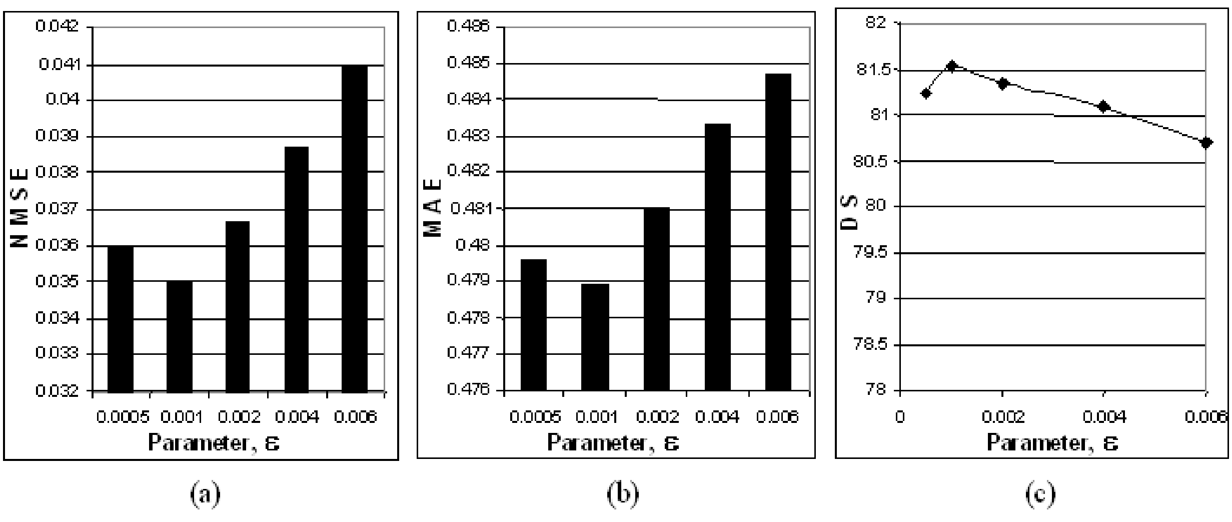

Fig. 3. Variation of NMSE, MAE and DS with parameter ϵ for predicting Japanese Yen exchange rate.

model did not perform inferiorly to ANN model.

In literature a number of techniques have been proposed to optimally determine the parameters C and ϵ. Mettera *et al.* suggested to set parameter C equal to the range of output values [23]. But its does not take into account the possible effect of outliers in training data. Hastie *et al.* proposed to estimate the optimal C value by cross validation on the data set [11] but this approach is computationally very expensive and data intensive. Kwok proposed asymptotically optimal values which are proportional to noise variance [20]. But this proposal does not reflect sample size while it is generally known that ϵ should be smaller for larger sample size. Hastie *et al.* proposed an approach where ϵ is optimally tuned for a particular noise density [11]. As Cherkassy *et al.* pointed out there have been many conflicting strategies suggested for setting the value of C and ϵ parameters [4]. Very recently, they proposed a methodology to analytically determine the value of C and ϵ. The method is claimed to determine "practical" values of C and ϵ, however, whether this method can select the best or optimum value is yet to be verified in a wide range of real world problems.

5. Conclusion

Earlier studies [17] reported that an ANN based forex model trained with scaled conjugate gradient learning algorithm predicted currency exchange rate more accurately than standard backpropgation based learning algorithm, and both the models outperformed traditional ARIMA model. Superiority of ANN model over ARIMA has also been supported by other studies [32]. For further improvement in prediction accuracy, we investigated the suitability of support vector machine in forecasting of exchange rates of six currencies against Australian dollar. An SVM based forecasting model was found to outperform the best ANN model measured on five commonly used evaluation metrics. Moreover, the performance of ANN model depends on a number of factors, e.g., the large number of degrees of freedom related to network architecture, tuning of learning parameters, complexity of the model etc. ANN training may stuck in local minima resulting in a very slow convergence and may require many trials before finding an optimal architecture that yields acceptably good performance.

In SVM, the only free parameters are C and ϵ. In this study, we observed that over a small range of value, C affected generalization performance moderately, but its effect was negligible when C exceeded certain threshold. This means that over a wide range of value, C has very little impact on prediction performance. The value of ϵ slightly affected the prediction accuracy. However, it did not cause any drastic degradation in performance. Future research direction will focus on achieving further improvement in prediction by considering other variations of support vector machines, use of complex or mixed kernels and automatic selection of parameters.

(Manuscript received Dec. 8, 2003,
revised June 16, 2004)

References

(1) G.E.P. Box and G.M. Jenkins: "Time Series Analysis: Forecasting and Control", Holden-Day, San Francosco, CA, (1990)

(2) L. Cao and F. Tay: "Financial forecasting using support vector machines", *Neural Comput & Applic*, Vol.10, pp.184-192 (2001)

(3) P. Chang and C. Osler: "Methodical madness: Technical analysis and the irrationality of exchange rate forecasts", *Economic Journal*, Vol.10, pp.636-661 (1999)

(4) V. Cherkassky and Y. Ma: "Practical Selection of SVM parameters and noise estimation for SVM regression", *Neural Networks*, Vol.17, pp.113-126 (2004)

(5) V. Cherkassky and P. Mullier: Learning from Data, Concepts, Theory and Methods, John Wiley (1998)

(6) G. Deboeck: Trading on the Edge: Neural, Genetic and Fuzzy Systems for Chaotic Financial Markets, New York Wiley (1994)

(7) R. Gencay: "Linear, nonlinear and essential foreign exchange rate prediction with simple technical trading rules", *J. Int. Econ*, Vol.47, pp.91-107 (1999)

(8) T. Gestel et. al.: "Financial time series prediction using least squares support vector machines within the evidence framework", *IEEE trans. Neural Network*, Vol.12, No.4, pp.809-821 (2001)

(9) J. Gooijer and D. Zerom: "Kernel-based multistep-ahead prediction of the US short-term interest rate", *J. Forecasting*, Vol.19, pp.335-353 (2000)

(10) S. Gunn: "Support vector machines for classification and regression", ISIS Technical Report, University of Southampton, UK (1998)

(11) T. Hastie, R. Tibshirani, and J. Friedman: *The elements of statistical learning:* Data mining, inference and prediction, Berlin: Springer (2001)

(12) S. Haykin: Neural Networks - A comprehensive Foundation, Prentice Hall (1999)

(13) T. Hill, M. O'Connor, and W. Remus: "Neural network models for time series forecasts", *Management Science*, Vol.42, pp.1082-1092 (1996)

(14) H.B. Hwarng and H.T. Ang: "Simple neural network for ARMA(p,q) time series", *OMEGA: Int. Journal of Management Science*, Vol.29, pp.319-333 (2002)

(15) W.C. Jhee and J.K. Lee: "Performance of neural networks in managerial forecasting", *Intelligent Systems in Accounting, Finance and Management*, Vol.2, pp. 55-71 (1993)

(16) I. Kaastra and M. Boyd: "Designing a neural network for forecasting financial and economic time-series", *Neurocomputing*, Vol.10, pp.215-236 (1996)

(17) J. Kamruzzaman and R. Sarker: "Forecasting of currency exchange rate: a case study", Proc. IEEE Int. Conf. Neural Networks & Signal Processing (ICNNSP'03), pp.793-797, China (2003)

(18) J. kkamruzzaman and R. Sarker: "Comparing ANN Baesd Models with ARIMA for Prediction of Forex Rates", *ASOR Bulletin*, Vol.22, No.2, pp.2-11 (2003)

(19) B.D. Klein and D.F. Rossin: "Data quality in neural network models: effect of error rate and magnitude of error on predictive accuracy", *OMEGA: Int. Journal of Management Science*, Vol.27, pp.569-582 (1999)

(20) J.T. Kwok: "Linear dependency between and the input noise in - support vector regression", *Proc. International Conf. Artificial Neural Networks*, pp.405-410 (2001)

(21) B. LeBaron: "Technical trading rule profitability and foreign exchange intervention", *Journal of Int. Economics*, Vol.49, pp.124-143 (1999)

(22) M. Leung, A.-S. Chen, and H. Daouk: "Forecasting exchange rates using general regression neural networks", *Computers & Operations Research*, Vol.27, No.11-12, pp.1093-1110 (2000)

(23) D. Mattera and S. Haykin: "Support vector machines for dynamic reconstruction of a chaotic system", In B. Scholkopf, J. Burges, & A. Smola (Eds.), Advances in kernel methods: Support vector machine, Cambridge, MA: MIT Press, (1999)

(24) M.C. Medeiros, A. Veiga, and C.E. Pedreira: "Modeling exchange rates: smooth transitions, neural network and linear models", *IEEE Trans. Neural Networks*, Vol.12, No.4, (2001)

(25) R. Meese and A. Rose: "An empirical assessment of nonlinearities in models of exchange rate determination", *Rev. Econ. Studies*, Vol.58, pp.603-619 (1991)

(26) M. El Shazly and H. El Shazly: "Forecasting currency prices using a genetically evolved neural network architecture", *International Review of Financial Analysis*, Vol.8, No.1, pp.67-82 (1999)

(27) Z. Tang and P.A. Fishwich: "Back-Propagation neural nets as models for time series forecasting", *ORSA Journal on Computing*, Vol.5(4), pp.374-385 (1993)

(28) V. Vapnik: Statistical Learning Theory, New York: Wiley, (1998)

(29) J.H. Wang and J.Y. Leu: "Stock market trend prediction using ARIMA-based neural networks", *Proc. IEEE Int. Conf. Neural Networks*, Vol.4, pp.2160-2165 (1996)

(30) S. Wang: "An insight into the standard back-propagation neural network model for regression analysis", *OMEGA: Int. Journal of Management Science*, Vol.26, pp.133-140 (1998)

(31) J. Yao, Y. Li, and C.L. Tan: "Option price forecasting using neural networks", *OMEGA: Int. Journal of Management Science*, Vol.28, pp.455-466 (2000)

(32) J. Yao and C.L. Tan: "A case study on using neural networks to perform technical forecasting of forex", *Neurocomputing*, Vol.34, pp.79-98 (2000)

(33) S. Yaser and A. Atiya: "Introduction to Financial Forecasting", *Applied Intelligence*, Vol.6, pp.205-213 (1996)

(34) G. Zhang and M.Y. Hu: "Neural network forecasting of the British Pound/US Dollar exchange rate", *OMEGA: Int. Journal of Management Science*, Vol.26, pp.495-506 (1998)

Joarder kamruzzaman (Non-member) received a B.Sc and M.Sc. in Electrical Engineering from Bangladesh University of Engineering & Technology, Dhaka, Bangladesh in 1986 & 1989 respectively, and a PhD in Information System Engineering from Muroran Institute of technology in 1993. Currently he is a faculty member in the Faculty of Information Technology, Monash University, Australia. His research interest includes neural networks, fuzzy system, genetic algorithm, computer networks and bioinformatics etc. He has published more than 60 refereed papers in international journals and conference proceedings. He is a program committee member of a number of international conferences.

Ruhul A. Sarker (Non-member) received his Ph.D. in 1991 from DalTech, Dalhousie University, Halifax, Canada, and is currently a senior faculty member in the School of Information Technology and Electrical Engineering, University of New South Wales, ADFA Campus, Canberra, Australia. His main research interests are Evolutionary Optimization, Neural Networks and Applied Operations Research. He has recently edited four books and has published more than 90 refereed papers in international journals and conference proceedings. Dr Sarker is actively involved with a number of national and international conferences & workshop organizations in the capacity of chair / co-chair or program committee member. He has recently served as a technical co-chair for IEEE-CEC2003. He is also a member of task force for promoting evolutionary multi-objective optimization operated by IEEE Neural Network Society.