

# Hybridizing Exponential Smoothing and Neural Network for Financial Time Series Predication

Kin Keung Lai<sup>1,2</sup>, Lean Yu<sup>2,3</sup>, Shouyang Wang<sup>1,3</sup>, and Wei Huang<sup>4</sup>

<sup>1</sup> College of Business Administration, Hunan University, Changsha 410082, China

<sup>2</sup> Department of Management Sciences, City University of Hong Kong,  
Tat Chee Avenue, Kowloon, Hong Kong  
{mskklai, msyulean}@cityu.edu.hk

<sup>3</sup> Institute of Systems Science, Academy of Mathematics and Systems Science,  
Chinese Academy of Sciences, Beijing 100080, China  
{yulean, sywang}@amss.ac.cn

<sup>4</sup> School of Management, Huazhong University of Science and Technology,  
1037 Luoyu Road, Wuhan 430074, China

**Abstract.** In this study, a hybrid synergy model integrating exponential smoothing and neural network is proposed for financial time series prediction. The proposed model attempts to incorporate the linear characteristics of an exponential smoothing model and nonlinear patterns of neural network to create a “synergetic” model via the linear programming technique. For verification, two real-world financial time series are used for testing purpose.

## 1 Introduction

A challenging task in financial markets such as stock market and foreign exchange market is to predict the movement direction of financial markets so as to provide valuable decision information for investors. Thus, various kinds of forecasting methods have been developed by many researchers and business practitioners. Of the various forecasting models, the exponential smoothing model has been found to be one of the effective forecasting methods. Since Brown [1] began to use simple exponential smoothing to predict inventory demand, the exponential smoothing models have been widely used in business and finance [2-3]. For example, Gardner [2] introduced exponential smoothing methods into supply chain management for predicting demand, and achieved satisfactory results. Leung et al. [3] used an adaptive exponential smoothing model to predict Nikkei 225 indices, and achieved good results.

However, the exponential smoothing method is only a class of linear model and thus it can only capture linear feature of financial time series. But financial time series are often full of nonlinearity and irregularity. Furthermore, as the smoothing constant decreases exponentially, the disadvantage of the exponential smoothing model is that it gives simplistic models that only use several previous values to forecast the future. The exponential smoothing model is, therefore, unable to find subtle nonlinear patterns in the financial time series data. Obviously, the approximation of linear models to complex real-world problems is not always sufficient. Hence, it is necessary to consider other nonlinear methods to complement the exponential smoothing model.

Recently, artificial neural network (ANN) models have shown their nonlinear modeling capability in financial time series forecasting. Since Lapedes and Farker [4] used ANN to predict the chaotic time series, the ANN models are widely used in the time series forecasting. For example, Refenes et al. [5] applied multilayer forward network models to forecast foreign exchange prices and obtained good results. Although the ANN models achieve success in financial time series forecasting, they have some disadvantages. Since the real world is highly complex, there exist some linear and nonlinear patterns in the financial time series simultaneously. It is not sufficient to use only a nonlinear model for time series because the nonlinear model might miss some linear features of time series data. Furthermore, previous studies [6-7] are shown that using ANN to model linear problems may produce mixed results.

In such situations, it is necessary to hybridize the linear model and nonlinear model for financial time series forecasting. This is because the ANN model and exponential smoothing model are complementary. On one hand, the ANN model can find subtle nonlinear features hidden in the time series data, but may miss some linear patterns when forecasting. On the other hand, the exponential smoothing model can give good results in the linear patterns of time series, but cannot capture the nonlinear patterns, which might result in inaccurate forecasts. Motivated by the previous findings, this study proposes a hybrid synergy model to financial time series prediction integrating exponential smoothing and ANN via linear programming technique.

The exponential smoothing model rather than other linear models such as ARIMA is chosen as neural network model's complement for several reasons. First of all, the major advantage of exponential smoothing methods is that they are simple, intuitive, and easily understood. These methods have been found to be quite useful for short-term forecasting of large numbers of time series. At the same time, exponential smoothing techniques have also been found to be appropriate in such applications because of their simplicity. Second, the exponential smoothing model has less technical modeling complexity than the ARIMA model and thus makes it more popular in practice. As Lilien and Kotler [8] reported, exponential smoothing models have been widely used by approximately 13% of industry. Third, Mills [9] found little difference in forecast accuracy between exponential smoothing techniques and ARIMA models. In some examples, exponential smoothing models can even obtain better results than neural network model. Foster et al. [10] once argued that the exponential smoothing is superior to neural networks in forecasting yearly data. Generally, the exponential smoothing model is regarded as an inexpensive technique that gives forecasts that is "good enough" in a wide variety of applications.

The remainder of the study is organized as follows. Sections 2 and 3 provide some basic backgrounds about the exponential smoothing and neural network methods for financial time series forecasting. In Section 4, the hybrid methodology combining the exponential smoothing and neural network model via linear programming is introduced. For verification, two experiments are performed in Section 5. Finally, Section 6 concludes the paper.

## 2 The Exponential Smoothing Forecasting Model

In the application of the exponential smoothing model, there are three types of models that are widely used in different time series. Simple exponential smoothing (Type I) is

used when the time series has no trend. Double exponential smoothing (Type II) is an exponential smoothing method for handling a time series that displays a slowly changing linear trend. Two approaches are covered: one-parameter double exponential smoothing, which employs a single smoothing constant; and Holt-Winters' two-parameter double exponential smoothing, which employs two smoothing constants. The third is Winters' method (Type III), which is an exponential smoothing approach to predicting seasonal data. This method also contains two approaches: multiplicative Winters' method, which is appropriate for increasing seasonal variation; and additive Winters' method, which is appropriate for constant seasonal variation [1-2].

In financial time series, there is irregularity, randomness and no trend. These features show that the simple exponential smoothing method is suitable for financial time series forecasting for the specified time period. Therefore, only the type I of the exponential smoothing model is described in detail. For the type II & III, interested readers can be referred to [1-3] for more details.

Suppose that the time series  $y_1, y_2, \dots, y_n$  is described by the model

$$y_t = \beta_0 + \varepsilon_t \tag{1}$$

where  $\beta_0$  is the average of the time series and  $\varepsilon_t$  random error. Then the estimate  $S_t$  of  $\beta_0$  made in time  $t$  is given by the smoothing equation

$$S_t = \alpha y_t + (1 - \alpha) S_{t-1} \tag{2}$$

where  $\alpha$  is a smoothing constant between 0 and 1 and  $S_{t-1}$  is the estimate of  $\beta_0$  in  $t-1$ .

Thus a point forecast made in time  $t$  for  $y_{t+1}$  is

$$\hat{y}_{t+1} = S_t = \alpha y_t + (1 - \alpha) \hat{y}_t \tag{3}$$

From Equation (2), we have

$$S_{t-1} = \alpha y_{t-1} + (1 - \alpha) S_{t-2} \tag{4}$$

Substituting Equation (4) to Equation (2), then

$$S_t = \alpha y_t + (1 - \alpha)(\alpha y_{t-1} + (1 - \alpha) S_{t-2}) = \alpha y_t + \alpha(1 - \alpha) y_{t-1} + (1 - \alpha)^2 S_{t-2} \tag{5}$$

Similarly, substituting recursively for  $S_{t-2}, S_{t-3}, \dots, S_1$  and  $S_0$ , we obtain

$$S_t = \hat{y}_{t+1} = \alpha y_t + \alpha(1 - \alpha) y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \dots + \alpha(1 - \alpha)^{t-1} y_1 + (1 - \alpha)^t S_0 \tag{6}$$

Here we see that  $S_t$ , the estimate made in time  $t$  of the average  $\beta_0$  of the time series, can be expressed in term of observations  $y_t, y_{t-1}, \dots, y_1$  and the initial estimate  $S_0$ . The coefficients measuring the contributions of the observations  $y_t, y_{t-1}, \dots, y_1$  – that is,  $\alpha, \alpha(1 - \alpha), \dots, \alpha(1 - \alpha)^{t-1}$  – decrease exponentially with time. For this reason this method is referred as simple exponential smoothing.

In order to use Equation (6) to predict the time series, we need to determine the value of the smoothing constant  $\alpha$  and the initial estimate  $S_0$ . For the smoothing constant, the ordinary least square (OLS) can be used to determine  $\alpha$ ; while for the

initial value  $S_0$ , we can let  $S_0$  be equal to  $y_1$ , i.e.  $S_0 = y_1$ , or let  $S_0$  be equal to the simple arithmetic average of a few previous observations. For example,  $S_0 = (y_1+y_2+y_3)/3$ . Once determining  $\alpha$  and  $S_0$ , exponential smoothing model can be used for prediction.

The advantage of the exponential smoothing method is that it is capable of fitting the linear patterns of the time series well and easy to use. But the financial time series is often irregular and nonlinear, it is not sufficient to use exponential smoothing for financial time series modeling.

### 3 The Neural Network Forecasting Model

In this study, one of the widely used ANN models, the back-propagation neural network (BPNN) [11], is used for time series forecasting. The main reason is that some studies (e.g. [11-12]) have shown that the BPNN with an identity transfer function in the output unit and logistic functions in the middle-layer units can approximate any continuous function arbitrarily well given a sufficient amount of middle-layer units [12]. Yu et al. [13] have also found that BPNN has been one popular model that can approximate various nonlinearities in the data series.

Generally, the BPNN can be trained by the historical data of a time series in order to capture the non-linear characteristics of the specific time series. The model parameters (connection weights and node biases) will be adjusted iteratively by a process of minimizing the forecasting errors. For time series forecasting, according to the previous computation process the relationship between the output ( $y_t$ ) and the inputs ( $y_{t-1}$ ,  $y_{t-2}$ , ...,  $y_{t-p}$ ) has the following mathematical representation.

$$y_t = a_0 + \sum_{j=1}^q a_j f(w_{0j} + \sum_{i=1}^p w_{ij} y_{t-i}) + e_t \tag{7}$$

where  $a_j$  ( $j = 0, 1, 2, \dots, q$ ) is a bias on the  $j$ th unit, and  $w_{ij}$  ( $i = 0, 1, 2, \dots, p; j = 0, 1, 2, \dots, q$ ) is the connection weights between layers of the model,  $f(\bullet)$  is the transfer function of the hidden layer,  $p$  is the number of input nodes and  $q$  is the number of hidden nodes. Actually, the BPNN model in (11) performs a nonlinear functional mapping from the past observation ( $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ ). to the future value ( $y_t$ ), i.e.,

$$y_t = \varphi(y_{t-1}, y_{t-2}, \dots, y_{t-p}, w) + e_t \tag{8}$$

where  $w$  is a vector of all parameters and  $\varphi$  is a function determined by the network structure and connection weights. Thus, in some senses, the BPNN model is equivalent to a nonlinear autoregressive (NAR) model.

A major advantage of neural networks is their ability to provide flexible nonlinear mapping between inputs and outputs. They can capture the nonlinear characteristics of time series well. However, using BPNN to model linear problems may produce mixed results [6-7]. Therefore, we can conclude that the relationship between exponential smoothing and BPNN is complementary. To take full advantage of the individual strengths of two models, it is necessary to integrate the exponential smoothing and BPNN models, as mentioned earlier.

### 4 The Hybrid Forecasting Methodology

In real life, financial time series forecasting is far from simple due to high volatility, complexity, irregularity and noisy market environment. Furthermore, real-world time series are rarely pure linear or nonlinear. They often contain both linear and nonlinear patterns. If this is the case, there is no universal model that is suitable for all kinds of time series data. Both exponential smoothing models and BPNN models have achieved success in their own linear or nonlinear domains, but neither exponential smoothing nor BPNN can adequately model and predict time series since the linear models cannot deal with nonlinear relationships while the BPNN model alone is not able to handle both linear and nonlinear patterns equally well [6]. On the other hand, as previously mentioned, for time series forecasting the relationship between exponential smoothing and BPNN is complementary. Exponential smoothing is a class of linear models that can capture time series' linear characteristics, while BPNN models are a class of general function approximators capable of modeling nonlinearity and which can capture nonlinear patterns in time series. Hybridizing the two models may yield a robust method, and more satisfactory forecasting results may be obtained by incorporating an exponential smoothing model and a BPNN model. Therefore, we propose a hybrid methodology integrating the exponential smoothing and the BPNN for financial time series forecasting. Different from decomposition-hybrid principle described in [6], we adopt the "parliamentary" hybridization strategy [14] to create a synergetic model, as shown in Fig. 1.

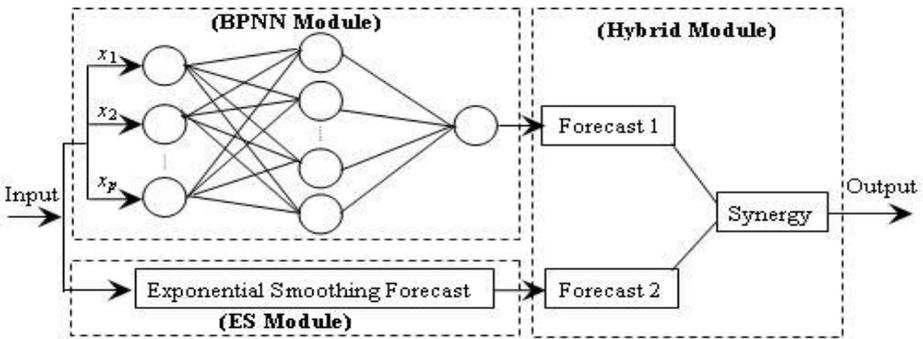


Fig. 1. The hybrid synergy forecasting model

From Fig. 1, the input is fed simultaneously into a BPNN model and an exponential smoothing forecasting model. The BPNN model generates a forecast result, while the exponential smoothing model also generates a time series forecast result. The two forecast results are entered into the hybrid forecast module and generate a synergetic forecast result as final output. In the hybridization process, the "parliamentary" hybridization strategy [14] is used, i.e.,

$$\hat{y}_t^{Hybrid} = \alpha \hat{y}_t^{ES} + (1 - \alpha) \hat{y}_t^{BPNN} \tag{9}$$

where  $\hat{y}_t^{ES}$  is the forecast result obtained from exponential smoothing model,  $\hat{y}_t^{BPNN}$  is the forecast result of the BPNN and  $\alpha$  is the weight parameter.

Through integrating linear patterns and nonlinear patterns of financial time series, a synergetic effect will be believed to be created to improve the predication performance. In Equation (9), a critical problem is how to determine the weight parameter  $\alpha$ . Generally, the value of  $\alpha$  can be estimated by the ordinary least square (OLS) method, i.e.,

$$MinQ = \sum_{t=1}^n (y_t - \hat{y}_t^{Hybird})^2 \tag{10}$$

However, its drawback of this approach is that the square treatment will move the fitted curve to some exceptional points and thus reducing the forecasting accuracy. One modification in this study is to minimize the sum of absolute error between estimated and the actual value, then

$$MinQ' = \sum_{t=1}^n |y_t - \hat{y}_t^{Hybird}| = \sum_{t=1}^n |e_t| \tag{11}$$

The Equation (11) can be solved by linear programming, let

$$u_t = \frac{|e_t| + e_t}{2} = \begin{cases} e_t, & e_t \geq 0 \\ 0, & e_t < 0 \end{cases}, \quad v_t = \frac{|e_t| - e_t}{2} = \begin{cases} 0, & e_t \geq 0 \\ -e_t, & e_t < 0 \end{cases} \tag{12}$$

Clearly,  $|e_t| = u_t + v_t, e_t = u_t - v_t$ , then the linear programming (LP) model can be formulated below.

$$(LP) \begin{cases} Min Q' = \sum_{t=1}^n (u_t + v_t) \\ \sum_{t=1}^n [y_t - (\alpha \hat{y}_t^{ES} + \beta \hat{y}_t^{BPNN})] - \sum_{t=1}^n (u_t - v_t) = 0 \\ \alpha + \beta = 1 \\ \alpha \geq 0, u_t \geq 0, v_t \geq 0, t = 1, 2, \dots, n \end{cases} \tag{13}$$

Using the simplex algorithm, an optimal hybridization parameter can be obtained from the LP problem in Equation (13). To verify the effectiveness of the hybridization approach, two experiments about exchange rate predication are performed.

### 5 Experiment Study

The data set used in this paper are daily data from 1 January 2000 till 31 December 2002 and are obtained from Pacific Exchange Rate Service (<http://fx.sauder.ubc.ca/>), provided by Professor Werner Antweiler, University of British Columbia, Vancouver, Canada. The daily data cover three years of observations of two major international currency exchange rates --- euros/US dollar (EUR/USD) and Japanese yen/US dollar (JPY/USD). In our empirical experiment, the data set is divided into two sample periods --- the estimation (in-sample) and the test (out-of-sample) periods. The estimation period covers observations from 1 January 2000 till 31 September 2002 and is used to estimate and refine the forecast model parameters. Meantime we take the data from 1 October 2002 to 31 December 2002 as test sets, which are used to evaluate the good or bad performance. For space, the original data are omitted, and detailed data can be obtained from the website.

In this study, the root mean square error ( $RMSE$ ) and directional statistics ( $D_{stat}$ ) [13] are used as evaluation criteria. In addition, the individual exponential and BPNN models are selected as benchmark models for comparison purposes. Finally, only one-step-ahead forecasting is considered in this study.

Based on our analysis above, we examine the forecast performances of two major currency exchange rates in terms of  $RMSE$  and  $D_{stat}$ . The experimental results are reported in Table 1.

**Table 1.** The experiment results of EUR/USD and JPY/USD

Currencies	EUR/USD			JPY/USD		
	ES	BPNN	Hybrid	ES	BPNN	Hybrid
$RMSE$	0.0047	0.0062	0.0035	0.6984	0.8226	0.6571
$D_{stat}$ (%)	56.45	62.90	67.74	51.61	58.06	66.13

From the viewpoint of  $RMSE$ , the hybrid methodology is the best for both EUR/USD and JPY/USD, followed by the individual exponential smoothing and individual BPNN model. For example of EUR/USD, the  $RMSE$  of the individual BPNN model is 0.0062, and the individual exponential smoothing model is 0.0047, while the  $RMSE$  of the hybrid methodology is only 0.0035.

However, the direction prediction is more important than the accuracy prediction in the financial markets because the former can provide decision information for investors directly. Furthermore, the high  $RMSE$  can not lead to high  $D_{stat}$ , as the exponential smoothing and the BPNN reveals. Focusing on the  $D_{stat}$ , the hybrid method performs the best, followed by the individual BPNN model; the worst is the individual exponential smoothing. For example of JPY/USD, the  $D_{stat}$  of exponential smoothing is only 51.61%, the  $D_{stat}$  of the BPNN model is 58.06%, while that of the hybrid methodology arrives at 66.13%. The main reason is that the hybrid methodology integrating linear patterns and nonlinear patterns creates a synergetic effect and thus improves the prediction performance.

## 6 Conclusions

In this paper, we propose a hybrid synergy methodology incorporating exponential smoothing and neural network for financial time series forecasting and explore the forecasting capability of the proposed hybrid synergy methodology from the point of level prediction and direction prediction. Experimental results obtained reveal that the hybrid methodology performs better than the two benchmark models, implying that the proposed hybrid synergy approach can be used as an alternative solution to the financial time series forecasting.

## Acknowledgements

This work is partially supported by National Natural Science Foundation of China (NSFC No. 70221001); Chinese Academy of Sciences; Key Research Institute of

Humanities and Social Sciences in Hubei Province-Research Center of Modern Information Management and Strategic Research Grant of City University of Hong Kong (SRG No. 7001677).

## References

1. Brown, R.G.: *Statistical Forecasting for Inventory Control*. New York, McGraw-Hill, 1959
2. Gardner, E.S.: Exponential Smoothing: The State of the Art. *Journal of Forecasting* 4 (1985) 1-28
3. Leung, M.T., Daouk, H., Chen, A.S.: Forecasting Stock Indices: A Comparison of Classification and Level Estimation Models. *International Journal of Forecasting* 16 (2000) 173-190
4. Lapedes, A., Farber, R.: *Nonlinear Signal Processing Using Neural Network Prediction and System Modeling*. Theoretical Division, Los Alamos National Laboratory, NM Report. No. LA-UR-87-2662, 1987
5. Refenes, A.N., Azema-Barac, M., Chen, L., Karoussos, S.A.: Currency Exchange Rate Prediction and Neural Network Design Strategies. *Neural Computing Applying* 1 (1993) 46-58
6. Zhang, G.P.: Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Neurocomputing* 50 (2003) 159-175
7. Denton J.W.: How Good are Neural Networks for Causal Forecasting? *Journal of Business Forecasting* 14 (1995) 17-20
8. Lilien, G.L., Kotler, P.: *Marketing Decision Making: A Model Building Approach*. New York, Harper and Row Publishers (1983)
9. Mills, T.C.: *Time Series Techniques for Economists*. Cambridge University Press (1990)
10. Foster, B., Collopy, F., Ungar, L.: Neural Network Forecasting of Short, Noisy Time Series. *Computers and Chemical Engineering* 16 (1992) 293-297
11. Hornik, K., Stinchcombe, M., White, H.: Multilayer Feedforward Networks are Universal Approximators. *Neural Networks* 2 (1989) 359-366
12. White, H.: Connectionist Nonparametric Regression: Multilayer Feedforward Networks can Learn Arbitrary Mappings. *Neural Networks* 3 (1990) 535-549
13. Yu, L., Wang, S.Y., Lai, K.K.: A Novel Nonlinear Ensemble Forecasting Model Incorporating GLAR and ANN for Foreign Exchange Rates. *Computers and Operations Research* 32 (2005) 2523-2541
14. Wedding II, D.K., Cios, K.J.: Time Series Forecasting by Combining RBF Networks, Certainty Factors, and the Box-Jenkins Model. *Neurocomputing* 10 (1996) 149-168