A Novel Modified Particle Swarm Optimization for Forecasting Financial Time Series

An-Pin Chen[1], Chien-Hsun Huang[1], Yu-Chia Hsu[1,2]

[1]Institute of Information Management, National Chiao Tung University, Hsinchu City, Taiwan
[2]Mackay Medicine, Nursing and Management College, Taipei City, Taiwan
apc@iim.nctu.edu.tw, katwin.huang@gmail.com, hyc0212@gmail.com

Abstract—Time series has been widely applied in the real world; traditional methods can hardly solve the dynamic environment issue resulting from the assumption of stationary process. Many traditional models and artificial intelligence technologies had been developed under this assumption, and adapted the dynamic environment based on the time-varying characteristic. But these models still has drawback of dividing the time series into training set and testing set when developing the models. It means the time-varying characteristic of these two sets did not be considered, and it might cause spurious regression phenomenon and result in misleading the statistic analysis. In order to forecast dynamic time series, a model which can consider the dynamic environment and conquer the out-of-sample problem is necessary. Particle swarm optimization (PSO) has the characteristics of fast-convergence and avoiding local optimal, also has been widely used in the time series forecasting. In this research, we proposed a modified PSO to consider the dynamic environment issue and use the advantage of PSO to forecast the dynamic financial time series.

Keywords- time series forecasting; particle swarm optimization; out-of-sample forecast

I. INTRODUCTION

In the real world, time series problem is everywhere. Financial and economic phenomenons, hydro meteorology, signal processing, even mechanical vibration can be solved by time series forecasting. In the research of time series, the knowledge of past experience and current situation are noticeable topic, especially the dynamic feature of time series always an issue that we need to pay attention to. In traditional research, there're GARCH and ARIMA model for solving these topics. For describing the volatility clustering phenomenon in stock price return volatility, ARCH model has been proposed and successfully integrated the simulation of dynamic environment [1]. GARCH model simplifies the parameters of ARCH model and improves the estimation performance [2]. On the other side, most of econometric models are constructed based on the assumption, time series is a stationary process. A stationary process is a stochastic process whose joint probability distribution does not change when shifted in time or space. It means the parameters such as the mean and variance, if they exist, do not change over time or position. In fact, many time series are non-stationary. If we process the non-stationary process as a stationary process, it will cause the temporary fluctuation has a long-term influence on whole process. Engle proves the analysis progressing the non-stationary process as a stationary process will lead to "spurious regression" phenomenon [3]. Spurious regression means there's no dependency relationship among variables, but make a wrong conclusion of the dependency relationship existing in the regression result. In certain situation, the certain linear combination of many non-stationary variables could form a stationary result, Engle called it "co-integration." In short, they applied the time-varying volatility feature in the proposed model and simulated the dynamic environment in a efficient way.

Recently, because of the searching ability, Neural Network (NN) and Genetic Algorithm (GA) has been widely applied in this aspect. By using the technique of sliding window, these technologies can adapt the dynamic environment. Along with the change of sliding window, the model can use the information to develop the model dynamically in the training set to forecast the time series. However, these models still could have the "spurious regression" phenomenon, it can adapt the dynamic environment in training phase.

Lately, Particle Swarm Optimization has been successfully applied to time series forecasting. Most of these researches use PSO in the parameter optimization to improve th search ability. Particle Swarm Optimization (PSO) is well-known as a population-based evolutionary algorithm, created by Eberhart and Kennedy [4]. The basic idea of PSO is a social behavior simulation for the bird flock and fish school. The behavior of all the individuals will be affected not only by the experience and cognition, but also the social behavior of whole society. Each single particle owns a speed and a direction separately in the hyperspace, and adjusts itself by proceeding probabilistic search strategy adjustment based on its own experience and the experiment of whole group. Consequently, PSO can improve the drawback of premature convergence and being trapped in the local optimum.

There is a limitation that PSO applied on the time series forecasting, if we use sliding window technique to adapt the dynamic environment, PSO will search the search space to find the best individual. But the knowledge in current sliding window cannot be used in the next forecasting process, because of the PSO in another sliding window will be counted as a different PSO. Conceptually, PSO seems to lie somewhere between genetic algorithms and evolutionary programming. [4] Like evolutionary programming, PSO is highly dependent on

stochastic processes. Like the crossover operation used in the genetic algorithm, the adjustment process of the particles move toward the best vertices is highly dependent on the individual and the group experience. Unique to the concept of particle swarm optimization is flying potential solutions through hyperspace, accelerating toward "better" solutions.[4] By referring the experience, the stochastic processes make sure that the evolution process will not just be replaced by the best answer we found so far, but ensure the exploration of unknown regions in the problem domain.

In the time series forecasting, PSO always assists neural network or other techniques to increase the diversity of the population and avoid being trapped in the local optimum.[5, 6]. We can hardly find other major application on the time series aspects. In our study, we proposed a modified particle swarm optimization applied on the dynamic time series problem. Combining the feature of fitness inheritance, sliding window, the threshold mechanism and PSO form a novel technology to forecasting the dynamic time series interval. The fitness inheritance feature make the algorithm can finely use the historical information; the sliding window can help the algorithm to adapt the dynamic environment; the threshold mechanism distinguish the false nearest neighbors to provide a better forecasting result. In our experiment, the modified PSO performs better than neural network. We hope to provide a novel method to solving the dynamic time series problem.

The remainder of this paper consists of five sections. Section 2 introduces the proposed model. Section 3 describes the experiment design. The forecasting performance of this method which proposed in this study is examined in Section 4. Then, the conclusion of this study is drawn in Section 6.

II. THE PROPOSED MODEL

We use the exchange rate as the example of time series. The deviation of exchange rates had been proved existing transition dynamics with macroeconomics fundamentals [7]. In order to forecast the exchange rate change, we have to understand the relationship among the exchange rate and all the factors could affect it. Obviously the standard PSO can hardly apply in this topic, because of the inability of processing time series. In this paper, we combine the inheritance feature of neural network and the evolution feature of particle swarm optimization.

In the modified PSO (m-PSO) we proposed, we define an evolution process of one particle set in a time unit as a run, and whole evolution process of all particle sets in a time unit as an iteration. We define the input variables as a data set DS, and the corresponding weights as a percentage set PS, which is a particle in the algorithm. The modification of PSO is being described in the following parts.

A. The Modified PSO

The m-PSO combines the features of fitness inheritance, sliding window and threshold mechanism to adapt the dynamic environment. In each iteration, the particles move based on the individual experience and the group experience toward the best vertices in the hyperspace. And the current iteration will inherits the experience of the previous iterations, and finely forecast the dynamic time series. The movement of the particles can be expressed by

$$PS_{i,j}^{new} = PS_{i,j} + V_{i,j}^{new} \qquad (1)$$

where

$PS_{i,j}^{new}$	The new position of ith percentage set in jth iteration
$PS_{i,j}$	The current position of ith percentage set in jth iteration
$V_{i,j}^{new}$	The new velocity of ith percentage set in jth iteration and is given by

$$V_{i,j}^{new} = wiV_{i,j} + c_1 r_{1,j}\left(pBest_j - PS_{i,j}\right)$$
$$+ c_2 r_{2,j}\left(gBest_j - PS_{i,j}\right) \qquad (2)$$

where

wi	The inertia weight that describes the degree of the particles' movement refer to the previously moved
$V_{i,j}$	The previous velocity of the ith percentage set in jth iteration
c_1, c_2	The local and global learning factors set at 2
$r_{1,j}$	the memory weight of the jth iteration, which describes the degree of the particles' movement refer to the best solution found so far in their experience
$r_{2,j}$	the cooperation weight of the jth iteration, which describes the degree of the particles' movement refer to the global best solution
$pBest$	the best solution found so far in current iteration
$gBest$	the global best solution of all particles found so far

The inertia weight   represents the exploration ability of the particle (or we can say percentage set). This parameter can grant the algorithm the exploration of unknown region discovery, and make sure not to converge in a local optimum. The consideration of actual dynamic environment could help us finely forecasting the time series.

B. Fitness Inheritance

From the past research [8-10], we can see the idea of fitness inheritance has been applied in many aspects, and successfully improve the model performance and lower the computation cost. In the fitness inheritance process, the accumulated knowledge can be used to forecast the time series in the future, and adjust the m-PSO to fit the dynamic environment. Along with the change of time, the m-PSO model will adopt the latest information to integrate the dynamic environment of the model.

It enhances the ability of considering dynamic environment, and mitigates the in-sample and out-of-sample problems affecting the time series interval forecasting process.

Figure 1 illustrated the fitness inheritance mechanism. In the beginning of an iteration, the initial vertices inherit the best vertices we found so far, excepts the first iteration. Each iteration can partially inherit the accumulated knowledge which produced by previous iterations and utilize it to progress the next particles' movement. Each iteration represents a period of time in the time series, the parameters will be optimized separately. In other words, the setting of parameters adjusts with the change of the dynamic environment. The consideration of best vertices has two type: one is local best vertex inheritance, another is global best vertex inheritance. From the angle of the certain iteration which pass through the accumulated knowledge to next iteration, the local best vertex inheritance means the best vertex in current iteration will be partially inherit to the next iteration; the global best vertex inheritance means the best vertex in all the iterations we found so far will be partially inherit to the next iteration. In the first iteration, the local best vertex and the global best vertex is the same. With the changing of the iterations, the scope of these two fitness inheritance types will be different. The characteristic of considering the group experience can be well-reflected the real situation.
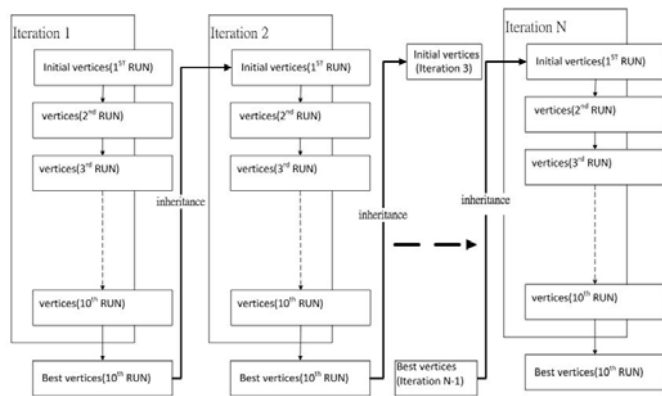

Figure 1. Fitness inheritance

C. Sliding Window

We add the sliding window to analysis and combine the information that provided in these days to form a conscientious decision process. In the whole experiment period, we can consider the subsample as a sample unit. The number of the subsamples will change along with the sliding window size w. As time goes by, the change of the nearest day will be considered and the change off the farthest day will be removed [11]. Meanwhile, we can see the change of input variables and try to figure out the phenomenon by observing the evolution process.

In this study, we attempt to take w days as a time unit for forecasting the change of exchange rate in next m days. By combining the change of input variables in a period of time, we expect this method can truly reflect the events happened in the real world. We used the "sliding window" shown in figure 2 to process the data and show. We reprocess the sample into N-$w+1$ subsamples as the input information and there will be N-w-m+1 output results to forecast the exchange rate change in next m days.
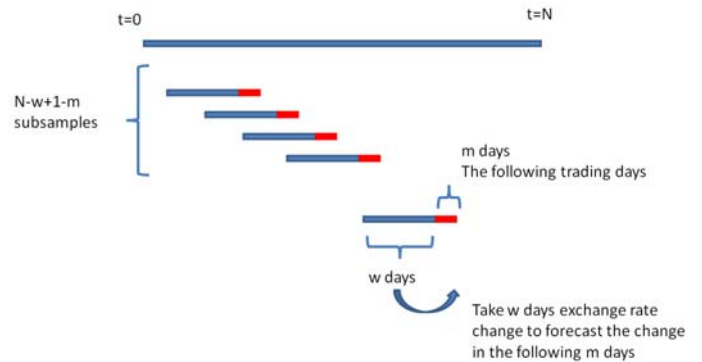

Figure 2. Sliding window

D. Threshold Mechanism

In the evaluation process, we add a threshold mechanism in evaluation process to perform the transition cost factor in real world, and eliminate the false nearest neighbor [12]. When an event that could affect the exchange rate happens, people always have the psychological expectancy and need to make a decision under the limitation of real world. This mechanism expresses the psychological state of the traders. As Hyginus and Serineh [13] said, proportional or "iceberg" costs create a band (thresholds) for the real rate, within which the marginal costs of arbitrage exceeds the marginal benefit. The reason of the proportional costs might induce to two aspects: One is sunk cost of international arbitrage and tendency for traders, who wait for sufficiently large arbitrage opportunities before making any moves, affects the increment of thresholds. Another point of view shows the evidence about the government intervention [14]. Because of the exchange rate could have a strong influence on the net export of a country, government might care about the large and persistent deviations as the cost of servicing debt denominated in foreign currency. But the detail of regime switching we do not discuss in this paper, this mechanism is focused on the phenomenon that results from regime switching.

We use trial-and-error method to decide the threshold in the experiment. By selecting the threshold from the ten percentages of the local best vertices, we can surely know the current situation in the evolution process from the dynamic threshold. If the difference between predict exchange rate change value ($PERC_i$) and actual exchange rate change value ($AERC_i$) is large, the current local best vertex (percentage set) will be judged as false near neighbors. It means the difference between $PERC_i$ and $AERC_i$ is beyond our expectation, and the fluctuation of exchange rate change is too large to control.

III. EXPERIMENT DESIGN

Focusing on the interaction among all variables which affects or be affected the exchange rate; we proposed this novel model to figure out the relationship. The spirit of this methodology is the process of combining the different effects

in many aspects, and we can conclude the suitable vertices in a natural way.

We collected 3957 daily exchange rate data of 14 country currencies gathering from the Central Bank of the Republic of China (Taiwan) in the period between 1993/1/5 and 2008/11/28. And, the monthly and daily economic indicators in the same period were gathered from the Taiwan Economic Journal. We divide all the data according to the time unit, each run represents the evolution process of data set in a time unit. In each iteration, we will try to find the best percentage set by the process of particle swarm. The experiment aimed to find the suitable weight set which can describe the exchange rate change well. According to 80/20 rule [15], there are 80 percentages of the data will be use as training data, others are the testing data. The initial percentage set of each iteration will inherit the result of previous iteration to control the bias of each evolution process.

We choose the input variables based on past research [16-18]. The daily exchange rate data of 13 main trading countries, including JPY/USD, GBP/USD, HKD/USD, KRW/USD, CAD/USD, SGD/USD, CNY/USD, AUD/USD, IDR/USD, THB/USD, MYR/USD, PHP/USD, EUR/USD, are chosen as the input variables. Other variables are the relative economic indicators of America, China, and Japan, which are the top 3 biggest trading partners of Taiwan, including the interest rate, import/export trade volume, and the consumer price index (CPI). The last variable is the 3-months treasury rate of USA. Total 23 variables are considered as the input variables for m-PSO to forecast the exchange rate of NTD/USD.

There are four stages of this experiment, the first stage is data pre-processing, the second one is initialize and the third, is evaluation, and last stage is the m-PSO.

In the first stage, we standardize the raw data to eliminate the influence which could be affected by unit of quantity. For instance, we collect the daily data, $DD_i = 1,2,3,...$, and we transform the daily data into a numeric volume data set, $DS = \{D_1, D_2, D_3,...,D_i\}$, without original unit of measure as follow:

$$D_i = \frac{DD_{i+1} - DD_i}{DD_i}, i = 1,2,3,...,n \qquad (3)$$

The DS includes all input variables change in a time unit, D_i means the change of ith data in a time unit.

In the second stage, we initialize a set of percentage, $PS = \{P_1, P_2, P_3,...,P_i\}$, which corresponds to data set DS and is a collection of variable percentage. Where P_i stands for the percentage of corresponding ith variable. The initial percentage set be randomly initialized based on normal distribution, the percentage which be initialized could be bigger than 1 or smaller than -1, because the percentage have all kinds of possibility. The initial percentage set could describe the relationship between input variable and exchange rate change in measurable criterion.

In the evaluation phase, there're so many possibilities that different combination of percentage set to forecast the exchange rate change, there's no deterministic answer.

Therefore, we try to provide a valuable result by means of the evaluation process. We calculate the fitness value FV by the following equation.

$$FV_j = \frac{1}{1 + \frac{AERC_i}{PERC_i}} - 0.5 \qquad (4)$$

If the predict exchange rate change is the same with the actual exchange rate change, and then the fitness value will be 0. It represents the percentage set can fully express the change of exchange rate.

In this experiment, we adopt the threshold mechanism to consider the transition cost of real world. We use the method that mentioned previously in this paper to decide the threshold adjusting itself by the situation. We think the fluctuation of exchange rate must reach a certain level and it will enhance the incentive of trading. Meanwhile, the suitable predict exchange rate change value is necessary for the trading decision, a percentage set which can provide a profitable prediction is indispensable.

Consequence, the accuracy of the m-PSO model is calculated based on the threshold mechanism. Only the forecasted percentage set reach a certain level which defined by the threshold is view as correct. Otherwise, the forecasting result is view as incorrect. The accuracy of the experiment is determined by Eq. (5).

$$\text{Accuracy} = \frac{\text{number of correct forecasting}}{\text{total number of forecasting}} \qquad (5)$$

In the last stage, the new position of percentage set will follow the equation (1). Each particle will take the new velocity to update its position in the search space. When moving the particles, the percentage of the corresponding data will change to enhance the prediction result. Not only adopting the inertia weight, we also consider the memory weight and the cooperation weight to discover the best combination of all information. In the current run, we expect to find the factors which significant affects the exchange rate and the better forecasting ability to predict the exchange rate change.

IV. EXPERIMENT RESULT

Several parameter settings of the experiment were performed in this study and listed in Table I. We separately discuss the experiment results in different conditions and try to find a rule of parameter setting in the experiment.

The experiment results of all the experiment in different number of particles under the condition of window size w=5 and forecasting exchange rate in m=3 trading days are plotted in figure 3. After learning the training, the m-PSO has accumulated a lot of knowledge to distinguish the dynamic environment and makes suggestions. Figure 4 shows the accuracy of m-PSO under different parameter settings from the aspect of sliding window size.

From the aspect of sliding window size, we show the accuracy under different sliding window size w, and there do

exist some differences. The accuracy curve with sliding window size w=40 has a significant gap with the accuracy curve with w=25. Despite the change of the particle number, the performance of the model with w=25 has the highest accuracy in forecasting exchange rate change.

And we can see the effect of changing particle number, the accuracy of the modified PSO performs an uptrend curve. After exceeding the 500 particle in the model, the curve mitigates the increment of the accuracy. Under certain particle number, there's positive correlation between the accuracy of the model and the number of the particle. The experiments with 500 and 1000 particles are significantly performs a higher accuracy. It indicates the particle number is an important factor to the model.

TABLE I.   THE PARAMETER SETTING OF MODIFIED PSO

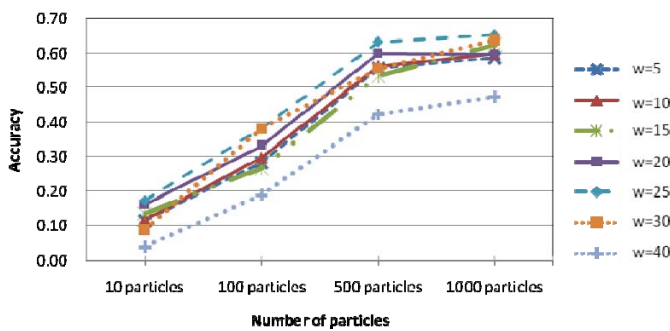| The sliding window size w	5,10,15,20,25,30,40
The forecasting trading day m	3
The number of particles	10,100,500,1000 particles
The number of runs	10 runs in one iteration

Figure 3.   The accuracy curve of modified PSO (m=3, training set) in the angle of different sliding window sizes
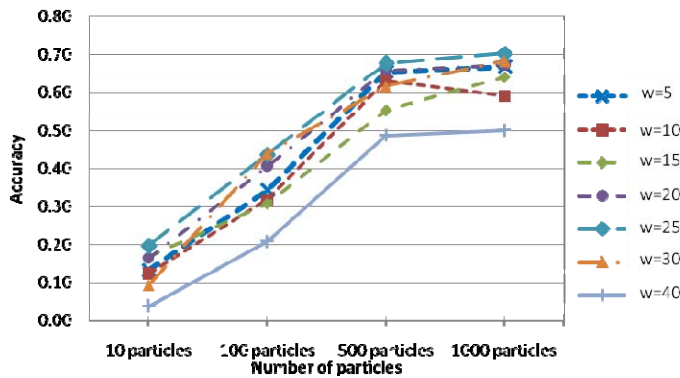

Figure 4.   The accuracy curve of modified PSO (m=3, testing set) in the angle of different sliding window size

V.   CONCLUSION

There might be deficient of considering the input variable about affecting the trend of financial time series in a certain period of time. This modified model enhances the original PSO model by combining the fitness inheritance, particle swarm optimization technique and the threshold mechanism, and mitigates the spurious regression phenomenon on dynamic time series forecasting. In this research, we also find out that there exist the optimal parameters to enhance the model performance and can be used for forecasting the exchange rate between NTD and USD.

REFERENCES

[1]   J.A. Frankel, K.A. Froot, 1987, Using survey data to test standard propositions regarding exchange rate expectations. American Economic Review, Vol. 77, pp. 133–153.

[2]   T. Bollerslev, 1986, Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, Vol.31, pp. 307-327.

[3]   D. Nelson, 1991, Conditional heteroskedasticity in asset returns: a new approach, Economics, Vol.8, pp. 347-370.

[4]   J. Kennedy, R. Eberhart.,1995, Particle Swarm Optimization, In Proceeding of IEEE International Conference on Neural Networks, Vol.4, pp. 1942-1948. .

[5]   X. Cai, N. Zhang, G. K.Venayagamoorthy, D. C. II Wunsch , 2007,Time Series Prediction with Recurrent Neural Networks Using a Hybrid PSO-EA Algorithm, In Proceeding of 2004 IEEE International Joint Conference on Neural Networks, Vol.2, pp. 1647-1652.

[6]   Y. Chen, B. Yang, J. Dong, 2006, Time-series prediction using a local linear wavelet neural network, Neurocomputing, Vol. 69, pp.449-465.

[7]   L. Kilian, M.P. Taylor, 2003, Why is it so difficult to beat the random walk forecast of exchange rates?, Journal of International Economics, Vol. 60, pp. 85-107.

[8]   J.H. Chen, D.E. Goldberg, S.Y. Ho, K. Sastry, 2002, Fitness Inheritance in Multi-Objective Optimization, In Proceeding of the Genetic and Evolutionary Computation, pp. 319-326.

[9]   M. Reyes-Sierra, C.A.C. Coello,2005, Fitness inheritance in multi-objective particle swarm optimization, In Proceeding of 2005 IEEE International Conference on Swarm Intelligence Symposium, pp. 116-123.

[10] M. Reyes-Sierra, C.A.C. Coello,2005, Improving PSO-Based Multi-objective Optimization Using Crowding, Mutation and -Dominance, In Evolutionary Multi-Criterion Optimization 2005, Lecture Notes in Computer Science 3410 , pp. 505–519.

[11] L Golab, M.T. Özsu, 2003, Issues in data stream management, In SIGMOD Record, Vol. 32, pp. 5-14.

[12] T. Edwards, D.S.W. Tansley, R. J. Frank, N. Davey, 1997, Traffic trends analysis using neural networks, In Proceeding of the International Workshop on Applications of Neural Networks to Telecommunications 3, Vol. 3, pp. 157–164.

[13] H. Leon. and S. Najarian, 2003,Asymmetric Adjustment and Nonlinear Dynamics in Real Exchange Rates, International Journal of Finance & Economics ,Vol. 10, pp. 15-39 .

[14] J. Dutta, H Leon, 2002, Dread of depreciation: Measuring real exchange rate interventions.

[15] R.L. Trueswell, 1969, Some Behavioral Patterns of Library Users: The 80/20 Rule,Wilson Libr Bull, Vol. 43, pp. 458-461.

[16] J.A. Frankel, K.A. Froot, 1987, Using survey data to test standard propositions regarding exchange rate expectations. American Economic Review, Vol. 77, pp. 133–153.

[17] S. Manzan, F.H. Westerhoff, 2007, Heterogeneous expectations, exchange rate dynamics and predictability, Journal of Economic Behavior & Organization, Vol. 64 , pp. 111–128.

N.C.Mark , 1995, Exchange rates and fundamentals: evidence on long-horizon predictability. American Economic Review 85, pp. 201–218.