Neural Fuzzy Forecasting of the China Yuan to US Dollar Exchange Rate — A Swarm Intelligence Approach

Chunshien Li[1], Chuan Wei Lin[1], and Hongming Huang[2]

Laboratory of Intelligent Systems and Applications
[1] Department of Information Management
[2] Department of Financial Economics
National Central University, Taiwan, R.O.C.
jamesli@mgt.ncu.edu.tw

Abstract. Exchange rate fluctuation has a significant effect on the risk of marketing business, economic development and financial stability. Accurate prediction for exchange rate may reduce commercial and economic risk arisen by exchange rate fluctuation. In this study, we propose an intelligent approach to the forecasting problem of the CNY-USD exchange rate, where a neuro-fuzzy self-organizing system is used as the intelligent predictor. For learning purpose, a novel hybrid learning method is devised for the intelligent predictor, where the well-known particle swarm optimization (PSO) algorithm and the recursive least squares estimator (RLSE) algorithm are involved. The proposed learning method is called the PSO-RLSE-PSO method. Experiments for time series forecasting of the CNY-USD exchange rate are conducted. For performance, the intelligent predictor is trained by several different methods. The experimental results show that the proposed approach has excellent forecasting performance.

Keywords: time series forecasting, neuro-fuzzy system (NFS), particle swarm optimization (PSO), recursive least-squares estimator (RLSE), hybrid learning, self-organization.

1 Introduction

Exchange rate systems can be divided into the categories of floating rate and fixed rate. Before the collapse of Bretton Wood System [1], banks adopted the fixed exchange rate system. After 1973, many countries have turned to adopt floating exchange rate. Since the floating exchange rate is changeably fluctuated, it may induce impact and risk on international trade, economic investment, and financial stability. Changes of exchange rate may come up with opportunity and risk. How to reduce the risk is a very interesting research topic. Many factors can affect exchange rates, for example, inflation rates, government intervention and economic growth [2]. The historical records in time sequential order of an exchange rate can be viewed as a time series [3]. Based on a prediction model, time series forecasting is to forecast future trend or change, based on past known observations.

Several approaches have been presented to predict exchange rate in literature [4]-[7]. Neural fuzzy system (NFS) approach [8] is one of the major methods. Many machine-

Y. Tan et al. (Eds.): ICSI 2011, Part I, LNCS 6728, pp. 616–625, 2011.

learning algorithms can be used to adjust the parameters of a NFS for application, such as, particle swarm optimization (PSO) [9] [10] [11], genetic algorithm (GA) [12], simulated annealing [13], simplex method [14], and downhill climbing method [15]. In this study, we use the theory of neuro-fuzzy system (NFS) to design an intelligent predictor for the time series forecasting problem of the CNY-USD exchange rate. In the NFS-based intelligent predictor, there are several fuzzy If-Then rules. The fuzzy rule can be divided into the Mamdani type and the Takagi-Sugeno (T-S) type [16] [17]. The difference between them lies on the Then-parts of the fuzzy rules. The Then-part of a T-S fuzzy rule is a function of inputs while a Mamdani fuzzy rule has a fuzzy set for its Then-part. Because T-S fuzzy rules possess greater capability of nonlinear functional mapping for applications than Mamdani fuzzy rules, we use T-S fuzzy rules for the NFS-based predicting approach in this study. For the formation of the NFS-based predictor in the study, there are two learning stages, the structure learning stage and the parameter learning stage. For the structure learning stage, we use a Fuzzy C-Means (FCM) based clustering method [18] to automatically determine the optimal number of fuzzy rules for the NFS-based predictor. Note that clusters generated the clustering method are regarded as fuzzy rules, based on the concept of input space partition. For the parameter learning stage, we develop a hybrid learning method, called the PSO-RLSE-PSO method, to fine-tune the parameters of the predictor. The hybrid learning method includes the well-known particle swarm optimization (PSO) algorithm and the recursive least squares estimator (RLSE) algorithm [16] [19]. In a hybrid way, the PSO algorithm updates the premise parameters of If-Then fuzzy rules, and the RLSE algorithm adjusts the consequent parameters. Afterward, the premise parameters are fixed, and the PSO algorithm is used again to update the consequent parameters, based on the result by the RLSE. With the PSO-RLSE-PSO hybrid learning algorithm, the intelligent can make better prediction accuracy.

In Section 2, the proposed methodology is specified, including the theory of NFS, the hybrid learning method and the FCM-based clustering algorithm for self-organization of the NFS-based predictor. In Section 3, experiments for the China Yuan (CNY) to US Dollar (USD) exchange rate forecasting by the proposed approach and the compared approaches are conducted, and the experimental results are given. Finally, a discussion is given is given and the paper is concluded.

2 Methodology of the Proposed Approach

2.1 Theory of Neuro-Fuzzy System (NFS)

The theory of fuzzy sets and fuzzy logic can be used to transform the experience of experts and knowledge into fuzzy If-Then rules, which are usually with uncertain information and imprecise expression. Fuzzy systems using this excellent property are usually known as an excellent problem-solving paradigm. Artificial neural networks have strong learning capability and they can adaptively learn to find hidden information and they have made great progress for extreme learning machines [20] [21]. Both of fuzzy systems and neural networks are universal approximator, by which any function can be approximated to any accuracy theoretically [22] [23].Integrating the advantages of fuzzy system and neural network, a NFS [10] [16] is a fuzzy inferential neural system, which can handle with complex issues. Suppose that there are K Takagi-Sugeno (T-S) type fuzzy rules [17] in a NFS. The ith fuzzy If-Then rule can be expressed as follows.

$$\begin{aligned} &\text{Rule } i: \quad \text{IF } x_1 \text{ is } s_1^i\big(h_1(t)\big) \text{ and ... and } x_M \text{ is } s_M^i\big(h_M(t)\big) \\ &\qquad\quad\ \ \text{Then } z^i = a_0^i + a_1^i h_1(t) + \cdots + a_M^i h_M(t) \end{aligned} \tag{1}$$

for i=1,2,…,K, where $\mathbf{H}(t)=[h_1(t)\ h_2(t)\ \ldots\ h_M(t)]^T$ is the crisp input vector to the NFS at time t, $\{x_i, i=1,2,\ldots,M\}$ are the input linguistic variables, $\{s_j^i, j=1,2,\ldots,M\}$ are the premise fuzzy sets of the ith fuzzy rule, z^i is the output, and $\{a_j^i, j=0,1,\ldots,M\}$ are the consequent parameters. Note that in the study the inputs are from historical data of exchange rate time series. The fuzzy inference of fuzzy system can be cast into neural-net structure with five layers to become the NFS [10] [16] [24] [25]. The explanation for the structure of NFS layer by layer is specified as follows. **Layer 1:** This layer is called the fuzzy-set layer. There are several nodes in the layer. Each node of the layer represents a fuzzy set in an input universe, by which the measured crisp input is transformed into fuzzy value. Each node output is a membership degree. In the study, fuzzy sets are designed using the Gaussian type membership function, given below.

$$\text{Gaussian}(h) = \exp\left(-\frac{1}{2}\left(\frac{h-m}{\sigma}\right)^2\right) \tag{2}$$

where h is a base variable, $\{m, \sigma\}$ are the mean and spread. **Layer 2:** This layer is called the firing-strength layer. The premise parts of fuzzy rules are formed in this layer. The nodes of the layer perform product operation (for t-norm operation) to calculate the firing strengths of the fuzzy rules, which are written as $\omega^i(t) = \prod_{j=1}^{M} s_j^i\left(h_j(t)\right)$, i=1,2,…,K. Note that $\{s_j^i, j=1,2,\ldots,M\}$ are the premise fuzzy sets of the i-th fuzzy rule specified in (1) and designed in (2) in the study. **Layer 3:** This layer is normalization layer. The nodes in the layer perform the process of normalization for the firing strengths of the fuzzy rules, given below.

$$\lambda^i(t) = \frac{\omega^i(t)}{\sum_{i=1}^{k} \omega^i(t)} \tag{3}$$

Layer 4: This layer is called the consequent layer. The nodes perform normalized consequents of all the fuzzy rules. Each node output represents a normalized consequent of fuzzy rule, given below.

$$\lambda^i(t) z^i(t) = \lambda^i(\mathrm{t})\left(a_0^{\mathrm{i}} + \sum_{j=1}^{M} a_j^i h_j(t)\right) \tag{4}$$

for i=1,2,…,K. **Layer 5:** This layer is called the output layer. The number of nodes is equal to that of the NFS outputs. In the study, we need one node only, which combines the node outputs of Layer 4 to produce the forecast for the CNY-USD exchange rate by the NFS output $\zeta(t)$ below.

$$\zeta(t) = \sum_{i=1}^{k} \lambda^i(t) z^i(t) = \sum_{i=1}^{k} \frac{\prod_{j=1}^{M} s_j^i\left(h_j(t)\right)}{\sum_{i=1}^{k} \prod_{j=1}^{M} s_j^i\left(h_j(t)\right)} \left(a_0^i + \sum_{j=1}^{M} a_j^i h_j(t) \right) \tag{5}$$

2.2 Clustering for Self-organization of the NFS Predictor

A Fuzzy C-means (FCM) based clustering method is used the self-organization of the intelligent NFS based predictor, by which the optimal number of fuzzy rules are automatically determined. The FCM algorithm, derived from K-means algorithm, was first proposed by Dunn and Bezdek [26]. Although FCM is better than K-means, but it still needs to enter a preset number of clusters for clustering. Haojun et al. improved this drawback and proposed the FCM-based splitting algorithm (FBSA) [18], which is based on the FCM method and a cluster validity index [27]. The FBSA can select the optimal number of clusters in between two preset positive integers, C_{min} and C_{max} which are the minimal and maximal numbers of clusters for the algorithm. The FBSA algorithm is given below. Step1: Preset C_{min} and C_{max}. Step 2: initialize C_{min} cluster centers ($\mathbf{V}$). Step 3: Do for-loop $c = C_{min}$ to C_{max}; 3.1: apply the FCM algorithm to update the membership matrix $\mathbf{U}$ and the cluster centers, $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_c\}$, where $\mathbf{v}_j = [v_{j,1}, v_{j,2}, \cdots, v_{j,Q}]^{\mathrm{T}}$ is the j-th cluster center; 3.2: test for convergence; if no, go to 3.1; 3.3: compute a validity value $V_d(c)$; 3.4: compute a score $S(i)$ for each cluster; split the worst cluster. Step 4: Compute the optimal number c_f such that the cluster validity index $V_d(c_f)$ is optimal. The validity index $V_d(.)$ is defined as follows.

$$V_d(\mathbf{U}, \mathbf{V}, c) = Scat(c) + \frac{Sep(c)}{Sep(C_{max})} \tag{6a}$$

$$Scat(c) = \frac{\frac{1}{c}\sum_{i=1}^{c} \|\sigma(\mathbf{v}_i)\|}{\|\boldsymbol{\sigma}(\mathbf{X})\|} \tag{6b}$$

$$Sep(c) = \frac{D_{max}^2}{D_{min}^2} \sum_{i=1}^{c} \left(\sum_{j=1}^{c} \left\| v_i - v_j \right\|^2 \right)^{-1} \tag{6c}$$

where $\boldsymbol{\sigma}(\mathbf{X}) = [\sigma(\mathbf{X})_1, \sigma(\mathbf{X})_2, \cdots, \sigma(\mathbf{X})_Q]^{\mathrm{T}}$, $\sigma(\mathbf{X})_p = \frac{1}{n}\sum_{k=1}^{n}(x_{k,p} - \bar{x}_{k,p})^2$, $D_{min} = min_{i \neq j}\left\|\mathbf{v}_i - \mathbf{v}_j\right\|$, $D_{max} = max_{i \neq j}\left\|\mathbf{v}_i - \mathbf{v}_j\right\|$ and n is the total number of data vectors in a given data set. Note that $Scat(c)$ represents the compactness among obtained clusters, whose range is between 0 and 1, and $Sep(c)$ represents the separation among the clusters.

2.3 Parameter Learning with PSO-RLSE-PSO Hybrid Learning Method

Particle Swarm Optimization (PSO) [9] is a swarm-based optimization algorithm, motivated by the food searching behavior by bird flocking or fish schooling in a society-oriented pattern. For a bird swarm, each bird is regarded as a particle whose location is viewed as a candidate solution to the problem. The search movement of the ith particle of the swarm is affected by two factors, the particle's best location in its experience and the swarm's best location during the search, denoted by $\mathbf{Pbest}_i$ and $\mathbf{Gbest}$, respectively. Assume that the search space is with Q dimensions. The location and velocity of the ith particle at time t are denoted as $\mathbf{X}_i(t)$ and $\boldsymbol{\psi}_i(t)$, respectively. The PSO algorithm is given by the following equations.

$$\boldsymbol{\psi}_i(t+1) = w \times \boldsymbol{\psi}_i(t) + c_1 \times \xi_1 \times \left(\mathbf{Pbest}_i - \mathbf{X}_i(t)\right) + c_2 \times \xi_2 \times (\mathbf{Gbest} - \mathbf{X}_i(t)) \tag{7a}$$

$$\mathbf{X}_i(t+1) = \mathbf{X}_i(t) + \boldsymbol{\psi}_i(t) \tag{7b}$$

where w in the inertia factor, $\{c_1,c_2\}$ are the learning rates, $\{\xi_1,\xi_2\}$ are random uniformly between 0 and 1. The procedure of PSO algorithm is given as follows. Step 1: initialize the position and velocity of particles. Step 2: calculate fitness for current position with fitness function $f(.)$. Step 3: update position and velocity of each particle. Step 4: If $f(\mathbf{Pbest}_i)$ gets improved, then update $\mathbf{Pbest}_i$. Step 5: If $f(\mathbf{Pbest}_i) < f(\mathbf{Gbest})$, then update $\mathbf{Gbest}$. Step 6: If termination condition is met, stop. Otherwise, go to Step 3 and the procedure continues.

In the general least squares estimation (LSE) problem, a linear model can be described as follows [19].

$$y = \theta_1 f_1(\mathbf{u}) + \theta_2 f_2(\mathbf{u}) + \cdots + \theta_n f_n(\mathbf{u}) + \varepsilon \tag{8}$$

where $\mathbf{u} = [u_1, u_2, \cdots, u_Q]^{\mathrm{T}}$ is the Q-dimensional input to the model, y is the observed data corresponding to the input $\mathbf{u}$, $\{f_1, f_2, \cdots, f_n\}$ are known functions of $\mathbf{u}$, $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_n]^{\mathrm{T}}$ are the parameter vector to be estimated, and ε is the model error. Note that the contents of the vector $\boldsymbol{\theta}$ can be viewed as the consequent parameters of the NFS predictor for the problem of time series forecasting. Observed data can be collected and used as training data for the proposed NFS predictor. The training data set with m data pairs is denoted as $\{(\mathbf{u}_j, y_j), j=1,2,\ldots,m\}$, where $(\mathbf{u}_j, y_j)$ is the j-th data pair in the form of (*input*, *target*). With the training data set, the LSE problem can be written in matrix form, given below.

$$\mathbf{Y} = \mathbf{A}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{9}$$

where $\mathbf{Y} = [y_1, y_2, \cdots, y_m]^{\mathrm{T}}$, $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m]^{\mathrm{T}}$ and $\mathbf{A}$ is the matrix formed by $\{f_i(\mathbf{u}_j), i = 1,2,\ldots,n \text{ and } j = 1,2,\ldots,m\}$. We denote the kth row of $\mathbf{A}$ and $\mathbf{Y}$ as $[\mathbf{a}_k^{\mathrm{T}}, y_k]$. Recursively, the optimal estimation for $\boldsymbol{\theta}$ can be obtained by the method of recursive least squares estimator (RLSE) [16], given below.

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \frac{\mathbf{P}_k \mathbf{a}_{k+1} \mathbf{a}_{k+1}^{\mathrm{T}} \mathbf{P}_k}{1 + \mathbf{a}_{k+1}^{\mathrm{T}} \mathbf{P}_k \mathbf{a}_{k+1}} \tag{10a}$$

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{P}_{k+1} \mathbf{a}_{k+1} (y_{k+1} - \mathbf{a}_{k+1}^{\mathrm{T}} \boldsymbol{\theta}_k) \tag{10b}$$

for $k=0,1,\ldots,m-1$. Initially, we set $\boldsymbol{\theta}_0$ to zero vector and $\mathbf{P}_0 = \alpha\mathbf{I}$, where α is a large positive number and $\mathbf{I}$ is the identity matrix. The optimal estimator for $\boldsymbol{\theta}$ is obtained after the last pair of training data $(\mathbf{u}_m, y_m)$ is done.

Table 1. Settings for PSO and RLSE

PSO		RLSE	
Swarm size	125	Consequent parameters	12
Initial particle position	Random in [0, 1]	$\boldsymbol{\theta}_0$	12×1 zero vector
Initial particle velocity	Random in [0, 1]	$\mathbf{P}_0$	$\alpha\mathbf{I}$, $\alpha = 10^{10}$
$\{w, c_1, c_2\}$	{0.8, 2, 2}	$\mathbf{I}$	Identity matrix

In the study, a hybrid learning method is devised for the parameter learning of the proposed NFS predictor, based on the result of self-organization by the FBSA clustering method to determine the optimal number of fuzzy rules for the intelligent predictor. There are two stages for the parameter learning. In **stage 1**, the PSO method is used to adapt the premise parameters of the NFS predictor and the RLSE method is use to update the consequent parameters. Both methods cooperate in hybrid way for fast learning purpose in the first stage. In **stage 2**, based on the result of the PSO-RLSE learning in the first stage, the PSO method is used again in the second stage to update the consequent parameters while the premise parameters are kept fixed. This two-stage learning method is called the PSO-RLSE-PSO method, for which the procedure is arranged below.

Step 1. Initialize necessary settings for the PSO and RLSE methods (in **stage 1**).
Step 2. Adapt the premise parameters of the NFS predictor by the PSO algorithm
Step 3. Update the consequent parameters by the RLSE algorithm.
Step 4. Calculate forecast error $e(t) = y(t) - \zeta(t)$, where $y(t)$ is the observed at time t and $\zeta(t)$ is the forecast by the NFS predictor.
Step 5. Calculate root mean squared error (RMSE), based on $\{e(t), t=1,2,\ldots m\}$, as follows.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{m}(e(t))^2}{m}} \tag{11}$$

Step 6. Update **Pbest** (for each particle) and **Gbest** for PSO.
Step 7. If termination condition for the PSO-RLSE learning (**stage 1**) is met, go to **Step 8** to proceed to **stage 2** of the parameter learning. Otherwise, go back to **Step 3**.
Step 8. Adapt the consequent parameters by the PSO (in **stage 2**). Calculate RMSE for all particles of the PSO. Update **Pbest** and **Gbest**.
Step 9. If termination condition for the PSO (**stage 2**) is met, then stop. Otherwise, go back to **Step 8** and the procedure continues.

3 CNY-USD Exchange Rate Forecasting by the Proposed Approach

The dataset recording the daily exchange rate of China Yuan (CNY) to US Dollar (USD) from 2006/01/01 to 2007/12/31 are obtained [28]. The dataset is normalized to [0, 1], from which the first 730 data are used as the training data and the rest 365 data are for testing purpose. For the proposed NFS predictor, the input vector is arranged as $\mathbf{H}(t)=[y(t\text{-}1),\ y(t\text{-}2)]^{\mathrm{T}}$ and the target is $y(t)$, where t is the time index. The FBSA clustering method is used for self-organization of the predictor to automatically determine the optimal number of fuzzy rules. Each cluster generated by the clustering method corresponds to a fuzzy If-Then rule. For the generated clusters, the cluster centers and their spreads in standard deviation are used for the initial settings of the premise fuzzy sets of the fuzzy rues. For parameter learning, the RMSE is used for the cost function. For the FBSA clustering method, the $C_{\min}$ and $C_{\max}$ are given as 2 and 10, respectively. After clustering, the curve of validity index versus cluster number is shown in Fig. 1, by which the optimal number of clusters is six. Thus, there are six rules generated automatically for the NFS predictor. The settings for PSO and RLSE are given in Table 1 and the parameters of the proposed PSO-RLSE-PSO are shown in Table 2. For the CYN-USD exchange rate forecasting, the NFS predictor with four

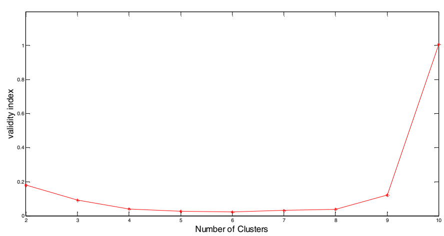

Fig. 1. Validity index versus cluster number by the FBSA clustering method

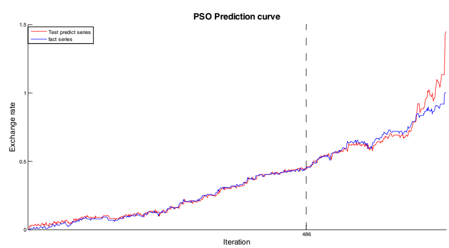

Fig. 2. Predicting response by the NFS predictor with six fuzzy rules (with PSO learning method only)

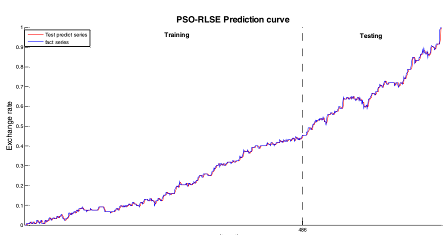

Fig. 3. Predicting response by the NFS predictor with six fuzzy rules (with PSO-RLSE learning method)

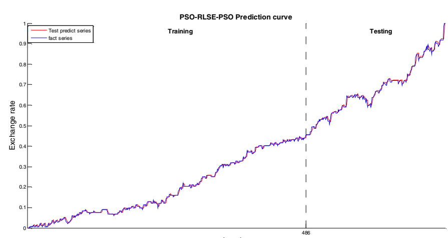

Fig. 4. Predicting response by the NFS predictor with six fuzzy rules (with PSO-RLSE-PSO learning method)

rules is trained by the PSO method only, the PSO-RLSE method, and the PSO-RLSE-PSO method, respectively. Ten trials are conducted for each experiment. The learning iterations for PSO and PSO-RLSE are set to 1000 and 500, respectively. The forecasting accuracy in RMSE for the training and testing are summarized in Table 2 for performance comparison.

Table 2. Parameters after learning by the PSO-RLSE-PSO

Premise-part = x_1 is $s_1(h_1(t))$ and x_2 is $s_2(h_2(t))$				
Parameter	m	σ	m	σ
Rule 1	0.7273	-6.5392	-3.006	-1.0228
Rule 2	2.32606	-9.0054	-4.2367	-1.5436
Rule 3	-9.3967	-3.1629	5.4079	3.5961
Rule 4	-3.4467	1.1636	3.3745	2.8554
Rule 5	-10.5639	2.5563	9.2401	0.3268
Rule 6	7.9862	1.6916	-10.4697	-3.3565

Consequent-part = $a_0+a_1h_1(t)+a_2h_2(t)$			
Parameter	a_0	a_1	a_2
Rule 1	-0.2478	-0.0592	0.9167
Rule 2	-0.0371	0.1809	0.6538
Rule 3	0.5980	0.1124	0.6040
Rule 4	0.2942	0.5874	0.5094
Rule 5	-9.0993e-06	-0.0438	-0.0010
Rule 6	0.9698	0.0017	0.0003

Table 3. Forecasting performance comparison in RMSE (mean/std)

Method	PSO	PSO-RLSE	PSO-RLSE-PSO
Training phase	0.1036/0.0696	0.0509/0.0484	0.0397/0.0126
Testing phase	0.1434/0.3048	0.0333/0.0316	0.0277/0.0022

Note that above results are based on ten trials.

4 Discussion and Conclusion

The self-organization neural fuzzy approach to the forecasting problem of the CNY-USD exchange rate has been presented. The design of the NFS-based predictor includes the structure-learning (self-organization) phase and the parameter learning phase. In the structure learning phase, there are no fuzzy rules at beginning. The NFS-based predictor is formed by the FBSA clustering method to automatically determine the optimal number of fuzzy If-Then rules and their complexity. After this, the parameter learning phase follows to fine-tune the predictor for good forecasting performance. There are three machine-learning methods used for parameter learning, Which are the PSO method, the PSO-RLSE method and the PSO-RLSE-PSO method. Although the PSO learning algorithm can adapt the predictor for the problem, the forecasting performance is merely plain. This could be because the parameter space was large, formed by the premise parts and the consequent parts, and the PSO might not easy to find the optimal or near-optimal solution. We further devised the PSO-RLSE hybrid learning method to train the NFS predictor and found the forecasting performance in accuracy got much improved, as shown in Table 3. The forecasting improvement may thank to the divide-and-conquer concept, based on which the PSO-RLSE method is devised. The parameter space was divided into two subspaces, the subspace of the premise parameters and the subspace of the consequent parameters. The PSO and the RLSE were used in hybrid way to search for the optimal solution, as stated previously. The PSO-RLSE method adapted the NFS predictor for forecasting not only in performance improvement but also in fast learning convergence. For this study, we further proposed the PSO-RLSE-PSO method to train the predictor to further improve the forecasting performance, and it worked, as shown in Table 3. The idea for the method is that based on the result by the PSO-RLSE, the PSO is used to further update the consequent parameters for better performance, while the premise parameters are fixed. Although the RLSE is good for linear regression model, it may not be good enough for the nonlinear forecasting problem, such as the CNY-USD exchange rate, which is usually fluctuated constantly and sometimes vibrated radically. This motivates the further use of PSO to the improvement search for the consequent parameters. Thus, the proposed PSO-RLSE-PSO method not only can preserve the merit of the PSO-RLSE method but also can deal with nonlinear forecasting problem with better performance.

The contribution of this research is three-fold. Firstly, the structure of the knowledge base in terms of fuzzy If-Then rules for the NFS-based intelligent predictor is automatically determined by the FBSA clustering method. Secondly, the innovation of the swarm-intelligence-based PSO-RLSE-PSO method is devised and applied successfully to the intelligent predictor for the nonlinear forecasting problem of the CNY-USD exchange rate. Thirdly, through performance comparison, excellent performance by the proposed approach has been shown.

Acknowledgment. This research work is supported by the National Science Council, Taiwan, ROC, under the Grant contract no. NSC99-2221-E-008-088.

References

1. Bordo, M.D., Eichengreen, B.J.: A retrospective on the Bretton Woods system: lessons for international monetary reform. University of Chicago Press, USA (1993)
2. Bodie, Z., Kane, A., Marcus, A.J.: Essentials of Investments. McGraw-Hill/Irwin, Boston (2004)
3. Hong, Y., Lee, T.H.: Inference on predictability of foreign exchange rates via generalized spectrum and nonlinear time series models. The Review of Economics and Statistics 85, 1048–1062 (2003)
4. Molodtsova, T., Papell, D.H.: Out-of-sample exchange rate predictability with Taylor rule fundamentals. Journal of International Economics 77, 167–180 (2009)
5. Leu, Y., Lee, C.P., Jou, Y.Z.: A distance-based fuzzy time series model for exchange rates forecasting. Expert Systems with Applications 36, 8107–8114 (2009)

6. Goldberg, E.: Optimal signal multi-resolution by genetic algorithms to support artificial neural networks for exchange-rate forecasting. Journal of Political Economy 104, 510–541 (1996)
7. Zhang, G., Hu, M.Y.: Neural network forecasting of the British pound/US dollar exchange rate. Omega 26, 495–506 (1998)
8. Jang, S.R.: ANFIS: adaptive-network-based fuzzy inference system. IEEE Transactions on System, Man, and Cybernetics 23, 665–685 (1993)
9. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: IEEE International Conference on Neuro Network Proceedings, vol. 4, pp. 1942–1948 (1995)
10. Chang, J.F., Chu, S.C., Roddick, J.F., Pan, J.S.: A parallel particle swarm optimization algorithm with communication strategies. Journal of Information Science and Engineering 21(4), 809–818 (2005)
11. Huang, H.C., Chen, Y.H., Abraham, A.: Optimized watermarking using swarm-based bacterial foraging. Journal of Information Hiding and Multimedia Signal Processing 1(1), 51–58 (2010)
12. Holland, J.H.: Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. University of Michigan Press, Ann Arbor (1975)
13. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by simulated annealing. Science 220, 671–680 (1983)
14. Nelder, J.A., Mead, R.: A simplex method for function minimization. The Computer Journal 7, 308–313 (1965)
15. Storey, C.: Applications of a hill climbing method of optimization. Chemical Engineering Science 17, 45–52 (1962)
16. Jang, J.S.R., Sun, C.T., Mizutani, E.: Neuro-Fuzzy and Soft Computing: a computational approach to learning and machine intelligence. Prentice Hall, USA (1997)
17. Xiu, Z.H., Ren, G.: Stability analysis and systematic design of Takagi-Sugeno fuzzy control systems. Fuzzy Sets and Systems 151, 119–138 (2005)
18. Sun, H., Wang, S., Jiang, Q.: FCM-based model selection algorithms for determining the number of clusters. Pattern Recognition 37, 2027–2037 (2004)
19. Goodwin, G.C., Sin, K.S.: Adaptive filtering prediction and control. Dover Publications, Inc., New York (2009)
20. Huang, G.B., Zhu, Q.Y., Siew, C.K.: Extreme learning machine: a new learning scheme of feedforward neural networks. In: Proceedings of International Joint Conference on Neural Networks (IJCNN 2004), Budapest, Hungary) (2004)
21. Huang, G.B., Liang, N.Y., Rong, H.J., Saratchandran, P., Sundararajan, N.: On-line sequential extreme learning machine. In: The IASTED International Conference on Computational Intelligence (CI 2005), Calgary, Canada (2005)
22. Nauck, D., Kruse, R.: Neuro-fuzzy systems for function approximation. Fuzzy Sets and Systems 101, 261–271 (1999)
23. Ying, H.: General interval type-2 Mamdani fuzzy systems are universal approximators. In: Fuzzy Information Processing Society, NAFIPS 2008, Annual Meeting of the North American, pp. 1–6 (2008)
24. Li, C., Lee, C.Y., Cheng, K.H.: Pseudo-error-based self-organizing neuro-fuzzy system. IEEE Transactions on Fuzzy Systems 12, 812–819 (2004)
25. Li, C., Lee, C.Y.: Self-organizing neuro-fuzzy system for control of unknown plants. IEEE Transactions on Fuzzy Systems 11, 135–150 (2003)
26. Dunn, J.C.: A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. Cybernetics and Systems 3, 32–57 (1973)
27. Ramze, R.M., Lelieveldt, B.P.F., Reiber, J.H.C.: A new cluster validity index for the fuzzy c-mean. Pattern Recognition Letters 19, 237–246 (1997)
28. OANDA, `http://www.oanda.com/currency/historical-rates`