2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

A Model with Fuzzy Granulation and Deep Belief Networks for Exchange Rate Forecasting

Ren Zhang, Furao Shen*, and Jinxi Zhao

***Abstract*— In recent years, neural networks is increasingly adopted in the prediction of exchange rate. However, most of them predict a specific number, which can not help the speculators too much because small gap between the predicted values and the actual values will lead to disastrous consequences. In our study, our purpose is to present a model to forecast the fluctuation range of the exchange rate by combining Fuzzy Granulation with Continuous-valued Deep Belief Networks (CDBN), and the concept of "Stop Loss" is introduced for making the environment of our profit strategy close to the real foreign exchange trade market. The proposed model is applied to forecasting both Euro/US dollar and British pound/US dollar exchange rate in our experiments. Experimental results show that the proposed method is more profitable in the trading process than other typical models.**

I. INTRODUCTION

ALTHOUGH speculative trade in foreign exchange market fuels exchange rate volatility to have large amplitude and high frequency, it increases the liquidity of the market and plays an indispensable role in stabilizing exchange market. Speculative trade is profitable, at the same time, the result from the changes in exchange rate are complex and diverse. This makes the exchange rate forecasting not only a very meaningful task, but also a challenging task which naturally attracts the attention of numerous researchers.

Early researchers developed classic statistical methods such as exponential smoothing approaches [1] [2] and autoregressive moving averages approach [3]. To deal with these nonlinear exchange rate data, quite a few improved statistical methods were further developed such as ARIMA [4] and ARCH [5]. As that neural networks (NNs) have general nonlinear function mapping capability which can approximate any continuous function to arbitrarily desired accuracy [6] [7], neural networks have been widely used in economic and financial fields since 1990s. White [8] is the first researcher to apply neural networks models in the financial fields. Kuan and Liu [9] use feedforward and recurrent neural networks to predict exchange rate. Even in recent years, neural networks are still popular for forecasting exchange rate. Fulcher et al. [10] apply Higher Order Neural Networks to forecast the AUD/USD exchange rate. Kiani and Kastens [11] forecast the GBP/USD, CAD/USD and JPY/USD exchange rates with feedforward and recurrent NNs. Chao et al. [12] forecast the GBP/USD and INR/USD exchange rates with Deep Belief Networks as they are able to find the global optimal solution.

Ren Zhang (zshank@163.com), Furao Shen (frshen@nju.edu.cn) and Jinxi Zhao (jxzhao@nju.edu.cn) are with the State Key Laboratory for Novel Software Technology at Nanjing University, Nanjing, 210046, P.R.China.
*Corresponding author is Furao Shen.

However, speculators employ leverage when they trade with others in foreign exchange market. Although the prediction errors of specific values are small, it will be magnified many times in the role of leverage. The magnified errors are intolerable to speculators. Rather than to reduce the prediction error, we predict the possible scope of exchange rate. A reliable range is more meaningful to making speculative decisions than a specific value which is not too reliable. The study of this problem is an important reference to speculators. In this paper, we focus on predicting the fluctuation range of exchange rate.

Granulation is one of the basic concepts that underlie human cognition [13]. Furthermore, Pedrycz declared that [14], "computational intelligence is a research endeavor aimed at conceptual and algorithmic integration of technologies of granular computing, neural networks and evolutionary computing". Hence, we granulate time series data of exchange rate. In this way, we predict the exchange rate whose volatility is caused by the combined effects of a variety of factors.

In this paper, we propose the CDBN-FG predictor which combines granulation with deep belief networks for continuous data [12] (CDBN) to forecast fluctuation range of exchange rate. Fuzzy information granulation (FIG) and CDBN which form the basis of CDBN-FG predictor is briefly introduced in section 2. In section 3, exchange rate series are granulated into granules Low, Mid and Up using five granulation methods described in [15], the possible minimums, medians and maximums in next granulation windows would be predicted by the CDBNs which are trained by the corresponding previous granules. Meanwhile, it is more important for speculators to transform the theoretical studies into applications in practical situations, we propose a trade strategy which applies the finance concept called "Stop Loss". Computational experiments are conducted to demonstrate the effectiveness and profitability of the proposed method in section 4.

II. THEORETICAL FOUNDATION

A. Fuzzy Information Granulation

Information granularity was firstly proposed by Zadeh [18] who gave a definition of information granulation with a general form: For a given universe of discourse U

$$f_{A \in U}(x) = \lambda \tag{1}$$

where x is a variable in U, A denotes one convex fuzzy subset of U, and λ represents the probability of the value

x belongs to the subset A. The concept of information granularity is employed in the fields such as information retrieval [16] and attribute reduction [17]. At different levels of granules, different information could be obtained, and the ordered granules often deliver more information than the unordered ones [15].

Information granules are information entities that emerge in the process of abstracting data and derivating knowledge from information [19]. The aim of information granulation is to decompose a given problems into some simple problems, while moving some redundant or irrelevant information. Approaches for information granulation include intervals [20] [21], rough sets [22]–[24], and fuzzy sets [18] [25]–[27]. These approaches can be generalized to two main categories: crisp information granulation (Crisp IG) and fuzzy information granulation (Fuzzy IG).

Crisp IG does not reflect the fact that in almost all of human reasoning and concept formation the granules are fuzzy [13]. In the face of such a complex foreign exchange market in our study, we adopt fuzzy IG. In this paper, we will apply five methods [15] including fuzzy sets and clustering methods to granulate time series before training and testing CDBNs. Two common membership functions are firstly reviewed here.

Let f be a mapping from a set A to $[0,1]$, that is, $f : A \to [0,1]$ and $f(x)$ denote the degree of the element x belonging to one fuzzy set on A.

(1) Triangular membership function
We can describe the triangular membership function (Trimf) in which there are three scalar parameters a, b, and c $(a < b < c)$ by the following formulation:

$$f(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \leq x \leq b \\ \dfrac{c-x}{c-b} & b \leq x \leq c \\ 0 & x > c \end{cases} \tag{2}$$

(2) Trapezoidal membership function
Trapezoidal membership function (Trapmf) is represented by four scalar parameters a, b, c, and d $(a < b < c < d)$, and given by

$$f(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \leq x \leq b \\ 1 & b \leq x \leq c \\ \dfrac{d-x}{d-c} & c \leq x \leq d \\ 0 & x > d \end{cases} \tag{3}$$

B. Deep Belief Networks for Continuous Data

Hinton et al. [28] introduced a greedy layer-wise unsupervised learning algorithm. Its basic idea is employing restricted Boltzmann machines (RBMs) as building blocks to construct a more powerful network named DBN. Based on the disadvantage that RBM can not deal with continuous values, Chao et al. [12] proposed a deep belief network for continuous data which trains a sequence of continuous restricted Boltzmann machines (CRBMs), and it was applied in forecasting exchange rate. We tentatively call it continuous-valued deep belief networks (CDBN) in which CRBMs are used instead of RBMs. A.F. Murray [29] employed RBM to model continuous data, however, this kind of RBM tends to generate continuous data with high symmetry. With continuous-valued stochastic units, CRBM offers improved ability with real continuous data. A CRBM is utilized to deal with continuous data as exchange rate in this study. A CRBM is given by Chen and Murray [30] with a simple and reliable training algorithm.

Let s_j be the output of neuron j , with inputs from neurons with states $\{s_i\}$.

$$s_j = \varphi_j\left(\sum_i w_{ij}s_i + \sigma \cdot N_j(0,1)\right) \tag{4}$$

with

$$\varphi_j(x_j) = \theta_L + (\theta_H - \theta_L) \cdot \frac{1}{1+e^{-a_j x_j}} \tag{5}$$

where σ is a constant. $N_j(0,1)$ represents a Gaussian random variable with zero mean and unit variance. $\varphi_j(x)$ is a sigmoid function with asymptotes at θ_L and θ_H. Parameter a_j is a noise-control parameter. It controls the slope of the sigmoid function, and thus the nature and extent of the units stochastic behavior [16]. Such behavior is similar to the noisy unit in [17]. The update equations for w_{ij} and a_j are

$$\Delta w_{ij} = \eta_w(< s_i s_j > - < s'_i s'_j >) \tag{6}$$

$$\Delta a_j = \frac{\eta_a}{a_j^2}(< s_j^2 > - < s'^2_j >) \tag{7}$$

where η_w and η_a are learning rates, s'_j denotes the one-step sampled state of unit j , and $< \cdot >$ refers to the mean over the training data. A simplified version of the same learning rule is used for the biases.

The training algorithm of a CDBN which is similar to that in [28] progresses on a layer-by-layer basis. First, a CRBM is trained directly on the input data. Hence, the neurons in the hidden layer of the CRBM can capture the essential features of the input data. The activations of the trained features are then used as "input data" to train the next CRBM. This training process continues until a prearranged number of hidden layers in the CDBN have been trained.

III. The CDBN-FG Predictors for Exchange Rate

In order to forecast the range of exchange rate volatility in a period of time, we propose a novel model which combines fuzzy granulation and CDBN, and call it CDBN-FG. Fig.1 gives the framework of CDBN-FG.

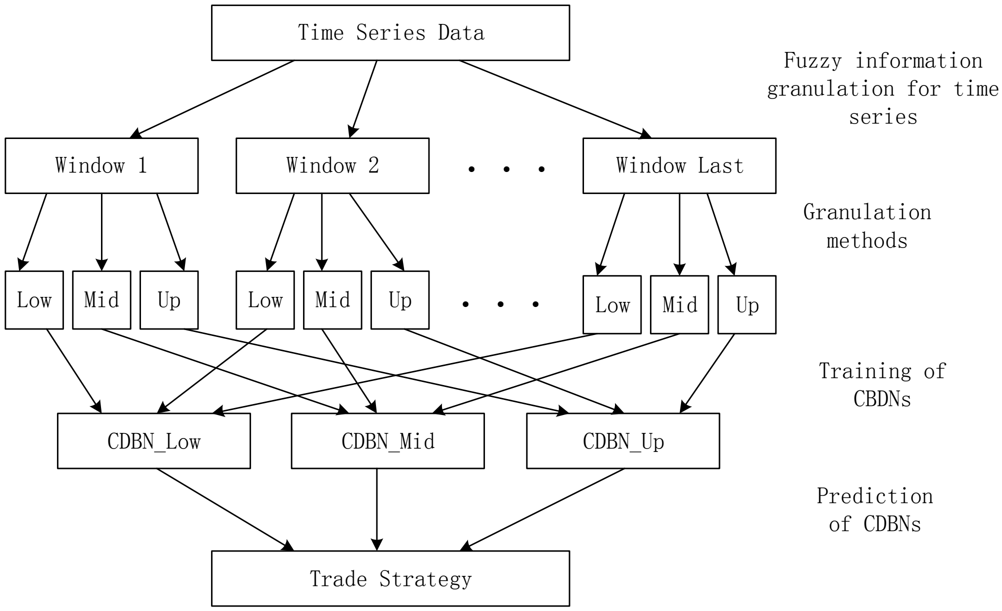

Fig. 1. The framework of CDBN-FG: 1) Split the original time series data to granulation windows which have the same length; 2) Granulate each window to three granules: Low, Mid and Up by granulation methods; 3) Train CDBNs: $CDBN_{Low}$, $CDBN_{Mid}$ and $CDBN_{Up}$ using the three granule series: Low, Mid and Up series; 4) Apply predictions of CDBNs into trade strategy.

A. Fuzzy Information Granulation for Time Series

The purpose of this paper is to propose a method for forecasting exchange rate which is a time series forecasting problem. The traditional way of forecasting exchange rate is to train and predict the $(n+1)$-th value using the first to the n-th values in the time series of exchange rate. However, in this study, the fluctuation range within a period which is represented by serval values together in time series will be predicted. In this case, the fuzzy granulation on time series will be employed.

Given that $Y = \{y_1, y_2, \ldots, y_n\}$ is a time series and w represents the length of granulation windows, $1 \leq w \leq n$. For the convenience to compare our proposed method to CDBN [12], we set the value w to 6 in our experiments, then the time series $\{y_1, y_2, \ldots, y_n\}$ can be split into $\{y_1, y_2, \ldots, y_w\}$, $\{y_{w+1}, y_{w+2}, \ldots, y_{w+w}\}$, $\ldots$, $\{y_{([n/w]-1)w+1}, y_{([n/w]-1)w+2}, \ldots, y_{([n/w]-1)w+w}\}$. $[n/w]$ subseries are produced, where $[x]$ represents the max integer not more than x.

How to granulate the values in each subseries with some fuzzy membership functions such as triangular and trapezoidal membership functions will be introduced in the next subsection.

B. Granulation Methods for the Values in Granulation Windows

Next, the granulation windows will be granulated into three granules: Low, Mid and Up, where the Low granule is granulated from the portion of each granulation window that has relatively small values, the Mid granule is produced from the middle values of this granulation window, and the Up granule is from the portion of this granulation window that has relatively lager values.

Given that the i-th subseries is $\{y_1^i, y_2^i, \ldots, y_w^i\}$, the granules low_i, mid_i and up_i are granulated by five granulation methods [15] which will be shown as follows.

1) Trimf-based granulation (trimf-Gr): Subseries $\{y_{1'}^i, y_{2'}^i, \ldots, y_{w'}^i\}$ is given to represent the sorted subseries of $\{y_1^i, y_2^i, \ldots, y_w^i\}$, then we divide it into $\{y_{1'}^i, y_{2'}^i, \ldots, y_{(w/2)'}^i\}$ and $\{y_{(w/2+1)'}^i, \ldots, y_{w'}^i\}$. The Low_i and Up_i granules are granulated from $\{y_{1'}^i, y_{2'}^i, \ldots, y_{(w/2)'}^i\}$ and $\{y_{(w/2+1)'}^i, \ldots, y_{w'}^i\}$ respectively according to the triangular membership function shown in Eq.(2). The Mid_i granule is represented by the median of the ordered subseries $\{y_{1'}^i, y_{2'}^i, \ldots, y_{w'}^i\}$. Then, the granulated granules based on Trimf are given by the following equations:

$$\begin{cases} low_i = \frac{2\sum_{j=1}^{w/2} y_{j'}^i}{w/2} - median\{y_{1'}^i, y_{2'}^i, \ldots, y_{w'}^i\} \\ mid_i = median\{y_{1'}^i, y_{2'}^i, \ldots, y_{w'}^i\} \\ up_i = \frac{2\sum_{j=w/2+1}^{w} y_{j'}^i}{w/2} - median\{y_{1'}^i, y_{2'}^i, \ldots, y_{w'}^i\} \end{cases} \tag{8}$$

where $median\{y_{1'}^i, y_{2'}^i, \ldots, y_{w'}^i\}$ represents the median of subseries $\{y_{1'}^i, y_{2'}^i, \ldots, y_{w'}^i\}$.

2) Trapmf-based granulation (trapmf-Gr): Being similar to Trimf-based granulation method, the trapezoidal membership function shown in Eq.(3) is used in trapmf-Gr method to granulate the values in each subseries. The subseries $\{y_1^i, y_2^i, \ldots, y_w^i\}$ is also sorted and divided into $\{y^i_{1'}, y^i_{2'}, \ldots, y^i_{(w/2)'}\}$ and $\{y^i_{(w/2+1)'}, \ldots, y^i_{w'}\}$, then granules Low_i and Up_i are also granulated from $\{y^i_{1'}, y^i_{2'}, \ldots, y^i_{(w/2)'}\}$ and $\{y^i_{(w/2+1)'}, \ldots, y^i_{w'}\}$ respectively. The Mid_i granule is also represented by the median of the ordered subseries $\{y^i_{1'}, y^i_{2'}, \ldots, y^i_{w'}\}$. The granulation formulates are:

$$\begin{cases} low_i = \frac{2\sum_{j=1}^{w/2} y^i_{j'}}{w/2} - y^i_{(w/2)'} \\ mid_i = median\{y^i_{1'}, y^i_{2'}, \ldots, y^i_{w'}\} \\ up_i = \frac{2\sum_{j=w/2+1}^{w} y^i_{j'}}{w/2} - y^i_{(w/2+1)'} \end{cases} \quad (9)$$

3) Minmax-based granulation (minmax-Gr): In minmax-based granulation method, we take the maximum in the subseries $\{y^i_1, y^i_2, \ldots, y^i_w\}$ as the Up_i granule and the minimum as the Low_i granule. The median of the subseries is used as the Mid_i granule. The minmax-Gr granulation formulates are:

$$\begin{cases} low_i = y^i_{1'} \\ mid_i = median\{y^i_{1'}, y^i_{2'}, \ldots, y^i_{w'}\} \\ up_i = y^i_{w'} \end{cases} \quad (10)$$

4) Hard c-means based granulation (HCM-Gr): The steps of using Hard c-means (HCM) clustering method to granulate the Low_i, Mid_i and Up_i granules from the subseries $\{y^i_1, y^i_2, \ldots, y^i_w\}$ are as follows:

Step 1: Randomly choose C data points from the subseries $\{y^i_1, y^i_2, \ldots, y^i_w\}$ as initial centers c_j, $j = 1, 2, \ldots, C$, $2 \le C \le w$.

Step 2: Cluster all data points in the subseries to C groups, in other words, to create a partition matrix $U = [u_{kj}]$, while the k-th data point in the subseries is the nearest one to the j-th center, $u_{kj} = 1$; otherwise, $u_{kj} = 0$. The distances from each data point to its nearest center is obtained by the standard Euclidean distance.

Step 3: Recalculate the new centers of each group $C_j = \sum_{u_{kj}=1} y^i_k / N$, where N is the number of data points in the j-th group.

Step 4: Calculate the objective function $J = \sum_{j=1}^{C}\sum_{k=1}^{w} \|y^i_k - c_j\|^2$. If $|J^t - J^{t-1}| < \xi$ where t is the iteration step and ξ is a given threshold, then stop; otherwise update the partition matrix U, and return to Step 3.

Here, we take $C = 3$ for three clusters: Low, Mid and Up. We set the smallest one as the Low_i granule, the middle one as the Mid_i granule, and the biggest one as the Up_i granule.

5) Fuzzy c-means based granulation (FCM-Gr): The steps of using Fuzzy c-means (FCM) clustering method to granulate the Low_i, Mid_i and Up_i granules from the subseries $\{y^i_1, y^i_2, \ldots, y^i_w\}$ are as follows:

Step 1: Randomly choose the value in $[0, 1]$ as u_{kj}, and according to to construct the initial partition matrix U with constraint that $\sum_{j=1}^{C} u_{kj} = 1, \forall k = 1, 2, \ldots, w$.

Step 2: Calculate the centers c_j using the formula $c_j = \sum_{k=1}^{w} u^m_{kj} y^i_k / \sum_{k=1}^{w} u^m_{kj}$ where m is an arbitrary real number greater than 1.

Step 3: Calculate the objective function $J = \sum_{j=1}^{C}\sum_{k=1}^{w} u^m_{kj}\|y^i_k - c_j\|^2$. If $|J^t - J^{t-1}| < \xi$ where t is the iteration step and ξ is a given threshold, then stop; otherwise update the partition matrix U by $u_{kj} = 1/(\sum_{p=1}^{C}(\|y^i_k - c_j\|/\|y^i_k - c_p\|)^{2/(m-1)})$, and return to Step 2.

Similarly, we also take $C = 3$ here. After getting the final three centers, we also set the smallest one as the Low_i granule, the middle one as the Mid_i granule, and the biggest one as the Up_i granule.

C. The Training and Prediction of CDBNs

In subsection B, three granules: Low, Mid and up are granulated in each granulation window. Suppose there are m windows segmented from a set of time series data, and we take the Low granule as an example to illustrate.

We extract out the Low granules from every granulation window, Low_i used to represent Low granule from the i-th granulation window. Then string them together to form a new time subseries $Low_1, Low_2, \ldots, Low_m$ for training $CDBN_{Low}$. The input data of $CDBN_{Low}$ are the past lagged observations of the Low granules subseries and the outputs are the predicted Low values. Each input sample is composed of a moving window of fixed length along the Low granules subseries.

Suppose $Low_1, Low_2, \ldots, Low_m$ are m granulation windows lagged Low granule observations of the exchange rate in the training set. We can use a network with p input nodes and one output node, if we need the one-step-ahead forecasts. The first training sample is composed of $\{Low_1, Low_2, \ldots, Low_p\}$ as the inputs and Low_{p+1} as the output. In the second training sample, $\{Low_2, Low_3, \ldots, Low_{p+1}\}$ are the inputs and Low_{p+2} is the target output. There are $m - p$ training samples in all. Training processes of $CDBN_{Mid}$ and $CDBN_{Up}$ are the same as that of $CDBN_{Low}$.

If the t-th window has been predicted, $\{Low_{t-p}, \ldots, Low_{t-1}\}$, $\{Mid_{t-p}, \ldots, Mid_{t-1}\}$ and $\{Up_{t-p}, \ldots, Up_{t-1}\}$ will be the input data of $CDBN_{Low}$, $CDBN_{Mid}$ and $CDBN_{Up}$ respectively, and Low_t, Mid_t and Up_t will be gotten.

D. Trade Strategy

How to utilize the predictions of forecasting exchange rate to obtain as many profits as possible? In traditional forecasting whose predictions are specific values, there is little reference information for the user. Because there are certain deviations between the predictions and the actual values, which will result in entirely different with that they want in the leveraged market. By contrast, there is a more practical significance in forecasting the exchange rate movements, that can guide us to open a long or a short order (a currency pair can be treated as a commodity, then the long order means buying this commodity and the short order means selling this commodity). In this study, we forecast not only whether exchange rate ascend or descend, but also the fluctuation range of exchange rate, which has great practical significance for speculator. In addition, the concept of “Stop Loss” in the field of financial transactions is introduced, which makes it possible that our theoretical research can be applied in real market.

As mentioned above, most of the transactions in foreign exchange market are leveraged transactions, thus only a slight volatility may result in great profits or losses. We can not guarantee that we are always correct in forecasting the market trends. When wrong predictions are made, we should bravely admit our mistakes with timely closing the wrong predicting orders to avoid more losses. For preventing the further loss, a stop loss price needs to be set at this time. We define the stop loss price and trends of prediction as follows:

$$Price_{SL} = \begin{cases} Mid - \beta(Mid - Low) & long\ order \\ Mid + \beta(Up - Mid) & short\ order \end{cases} \quad (11)$$

and

$$T_i = sign(P_i - R_{i-1}) = \begin{cases} 1 & up\ trend \\ -1 & down\ trend \end{cases} \quad (12)$$

where β is a parameter using to adjust the size of the stop loss interval. $P_i = (Low_i + Mid_i + Up_i)/3$ is the average of the predicting Low_i, Mid_i and Up_i granules on the i-th granulation window. R_{i-1} represents the close price of the $(i-1)$-th granulation window. T_i is the predictable trend of the i-th granulation window. The trading period is one day in our experiments. The trading strategy we propose could be described as follows: according to the predictable trend, we enter into the market with opening a long order or a short order at open price on every day (close price on the previous day), and then stop loss price is set on this day, where “open price” and “close price” are concepts in finance field to represent the start price and the end price in a time period. Because foreign exchange trading is 24 hours a day, we use the close price on the previous day as the open price on this day. If real-time price in the market pass through the stop loss price, we get out of the market at the price at that moment, otherwise at close price on this day.

We define the trade strategy of CDBN-FG as follows:

$$Profit = \sum_i T_i \cdot (Price_i - R_{i-1}) \quad (13)$$

where $Price_i$ is the price exiting the market at the i-th day, and

$$Price_i = \begin{cases} y^i_{\arg\min_k \ y^i_k \le Price^i_{SL}} & T_i = 1 \\ y^i_{\arg\min_k \ y^i_k \ge Price^i_{SL}} & T_i = -1 \end{cases} \quad (14)$$

where y^i_k is the k-th value in the i-th granulation window, $1 \le k \le w$. If $y^i_k \le Price^i_{SL}$ or $y^i_k \ge Price^i_{SL}$ cannot be satisfied then $k = w$, at this time, $Price_i = R_i$.

We compare CDBN-FG with CDBN [12] and BP-FG in experiments. The trade strategy of BP-FG is the same as that of CDBN-FG. The trade strategy of CDBN [12] is given by

$$Profit = \sum_i sign(P_i - R_{i-1})(R_i - R_{i-1}) \quad (15)$$

where P_i represents the i-th prediction.

IV. Experiments

A. Data

The original time series of CDBN-FG and BP-FG which are composed of the close prices of every 4 hours in the foreign exchange market are used in our experiments. In this case, a granulation window including 6 forex market data just represents one day. In our experiments, we use two data sets of exchange rates which are Euro/US dollar (EUR/USD) from Jan. 3, 2011 to Dec. 29, 2012 and British pound/US dollar (GBP/USD) from Sep. 28, 2009 to Oct. 28, 2011. The first data set consists of 761 observations of granulation windows totally, and we use 701 observations as training set and the remaining 60 observations to test. Similarly, we get a total 726 observations of granulation windows, 666 observations are kept for training and the other 60 observations are also remaining to use as test samples. The data which are used for training CDBN [12] consist of close price of each day. All these data are real historical data on the foreign exchange market. We download it from (`http://www.dukascopy.com/swiss/english/marketwatch/historical/`). Below, we use a simple method to make the two data sets be between 0 to 1.

$$y_t = p_t - 1 \quad (16)$$

where p_t represents the foreign exchange rate at time t.

B. Experiments Design

The purpose of this study is to forecast the approximate fluctuation range of exchange rate, and discuss whether there are considerable profits on the standard of realistic trading market using the proposed prediction model.

The two original data sets of EUR/USD and GBP/USD exchange rate series are normalized by formula (16). Then they are granulated to many granulation windows each of which span just a day by fixing the length of the window to 6. Three granules that Low, Mid and Up are granulated from each granulation window by the five granulation methods given in Section 3. Then three time subseries which are strung by the granules could build three CDBNs: $CDBN_{Low}$, $CDBN_{Mid}$ and $CDBN_{Up}$ respectively. At last, the trade strategy mentioned in the previous section are applied on the 60 testing samples in each data set. We use the criteria “how much profit” to evaluate the performance of the predictors. For the stability and repeatability of the experiments, we train each CDBN 100 times by using 100 sets of different initial values, and the averages of 100 runs are finally taken as the experimental results.

The parameters of our model are same as [12]: it has 6 input nodes, 16 and 8 hidden nodes in two hidden layers respectively, 1 output node, that is a 6-16-8-1 architecture. For the experiments' comparability, three BPs with 6-16-8-1 architecture are used in BP-FG.

C. Results

TABLE I represents the maximum profits based on CDBN-FG and CDBN [12] with two currency pairs. 'pip' is the unit of profit. 1 pip is 10,000 times to the price difference. For example, suppose we have 10,000 dollars, according to the general risk tolerance, we can buy or sell 1 lot commodities (currency pair). So, we earn 10 dollars per pip profit on 1 lot. In this case, we will earn 5,000 dollars if 500 pips profit is obtained in 60 days, that means our funds could be up to 50 percents, or we profit by 50% in two months. Obviously, the more pips means the greater profit. TABLE I shows that CDBN-FG which employs any one of the five granulation methods can get more profits than CDBN [12]. The greatest profit of CDBN-FG on EUR/USD is 743.819 pips, which is 548.87% more than 114.633 pips obtained by CDBN [12]. The greatest profit of CDBN-FG on GBP/USD is 366.856 pips, which is 30.67% more than 280.775 pips obtained by CDBN [12].

Fig.2 and Fig.3 represent the profits of CDBN-FG and BP-FG on EUR/USD and GBP/USD respectively. The horizontal axis in the figures are the values of parameter β with a variety of settings, and the vertical axis indicates the profit in the corresponding case.

Fig.2 and Fig.3 show that CDBN-FG can obtain much greater profits than BP-FG on both EUR/USD and GBP/USD. The lowest profit of CDBN-FG on EUR/USD is 324.109 pips and the greatest profit of BP-FG on EUR/USD is 224.548 pips. The lowest profit of CDBN-FG on GBP/USD is 203.613 pips and the greatest profit of BP-FG on GBP/USD is 91.709 pips. Thus, we can know that the lowest profits of CDBN-FG are at least 100 pips greater than the greatest profits BP-FG can obtain.

In order to verify the sensitivity of the parameter β in CDBN-FG, we try to set β from 0.5 to 10. Fig.2 and Fig.3 show: 1) While the greatest profits are gotten, most of β values are the intermediate values between 0.5 to 10. This indicates that the β value is not the larger the greater profit, or not the smaller the better. 2) While we change the β values which are greater than 5, the changes of the corresponding profits are small.

We can learn from Fig.2 and Fig.3 that different granulation methods suit different currency pairs according to their attribute difference. Trimf-based granulation method is used to obtain the greatest profit on EUR/USD, while trapmf-based granulation method is used to obtain the greatest profit on GBP/USD.

V. Conclusion

There is an increasing interest in forecasting foreign exchange rate recently. In this paper, we propose a model combining Fuzzy Granulation and Continuous-valued Deep Belief Networks to forecast the fluctuation range of exchange rate. Both daily EUR/USD and GBP/USD exchange rate are used in this study. We compare the proposed CDBN-FG with CDBN [12] and BP-FG on these data sets.

By using the criteria "how much profit can it obtain" to evaluate the predictors, we can draw a conclusion that the proposed model CDBN-FG in this paper is more profitable than both a classical approach and a newly proposed efficient neural networks in forecasting foreign exchange rate.

Acknowledgements

This work is supported in part by the National Science Foundation of China under Grant Nos. (61375064, 61373001, 61321491), 973 program (2010CB327903), and Jiangsu NSF grant (BK20131279, BK2011567).

References

[1] J.F. Muth, "Optimal properties of exponentially weighted forecasts", *Journal of the American Statistical Association.*, vol. 290, no. 55, pp. 299-306, 1960.
[2] C.C Pegels, "Exponential smoothing: some new variations", *Management Science*, vol. 15, no. 5, pp. 311-315, 1965.
[3] H. Markowitz, "Portfolio selection", *The journal of finance.*, vol. 7, no. 1, pp. 77-91, 1952.
[4] W.Enders, "ARIMA and cointegration tests PPP under fixed and flexible exchange rate regimes", *The review of economics and statistics*, vol. 70, no. 3, pp. 504-508, 1988.
[5] F.X. Diebold and M. Nerlove, "The dynamics of exchange rate volatility: A multivariate latent factor ARCH model", *Journal of applied econometrics*, vol. 4, no. 1, pp. 1-21, 1989.
[6] G. Cybenko, "Approximation by superposfitions of a sigmoidal function", *Mathematical Control Signals*, vol. 2, pp. 303-314, 1989.
[7] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators", *Neural Networks*, vol. 2, pp. 359-366, 1989.
[8] H. White, "Economic prediction using neural networks: The case of IBM daily stock returns", *in Proc. 2nd Annu. IEEE Conf. on Neural Networks*, San Diego, pp. 451-458, 1988.
[9] C. M. Kuan and T. Liu, "Forecasting exchange rates using feedforward and recurrent neural networks", *Journal of Applied Econometrics*, vol. 10, pp. 347-364, 1995.
[10] J. Fulcher, M. Zhang and S. Xu, "The application of higher-order neural networks to financial time series prediction", *Artificial Neural Networks in Finance and Manufacturing*, Idea Group Publishing, Hershey, PA London, 2006.
[11] K. Kiani and T. Kastens, "Testing forecast accuracy of foreign exchange rates: predictions from feedforward and various recurrent neural network architectures", *Computational Economics*, vol. 32, no. 4, pp. 383-406, 2008.
[12] J. Chao, F. Shen and J. zhao, "Forecasting Exchange Rate with Deep Belief Networks", *International joint conference on neural networks*, San Jose, California, USA, pp. 1259-1266, 2011.
[13] L.A. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic", *Fuzzy Sets and Systems*, vol. 90, pp. 111-127, 1997.
[14] W. Pedrycz, "Computational Intelligence: An Introduction", CRC Press, 1997.
[15] J. Ruan, X. Wang and Y. Shi, "Developing fast predictors for large-scale time series using fuzzy granular support vector machines", *Applied soft computing*, vol. 13, pp. 3981-4000, 2013.
[16] J.T. Yao and Y.Y. Yao, "Information granulation for Web based information retrieval Support Systems", *Data Mining and Knowledge Discovery*, vol. 138, 2003.
[17] Q. Hu, Z. Xie and D. Yu, "Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation", *Pattern Recognition*, vol. 40, no. 12, pp. 3509-3521, 2007.
granulation

TABLE I

THE MAXIMUM PROFITS BASED TWO MODELS ON TWO CURRENCY PAIRS RESPECTIVELY

Profit	CDBN-FG					CDBN [12]
	minmax-Gr	HCM-Gr	FCM-Gr	trimf-Gr	trapmf-Gr	
EUR/USD	723.33	490.051	613.954	**743.819**	628.167	114.633
GBP/USD	355.31	295.405	334.593	318.751	**366.856**	280.775

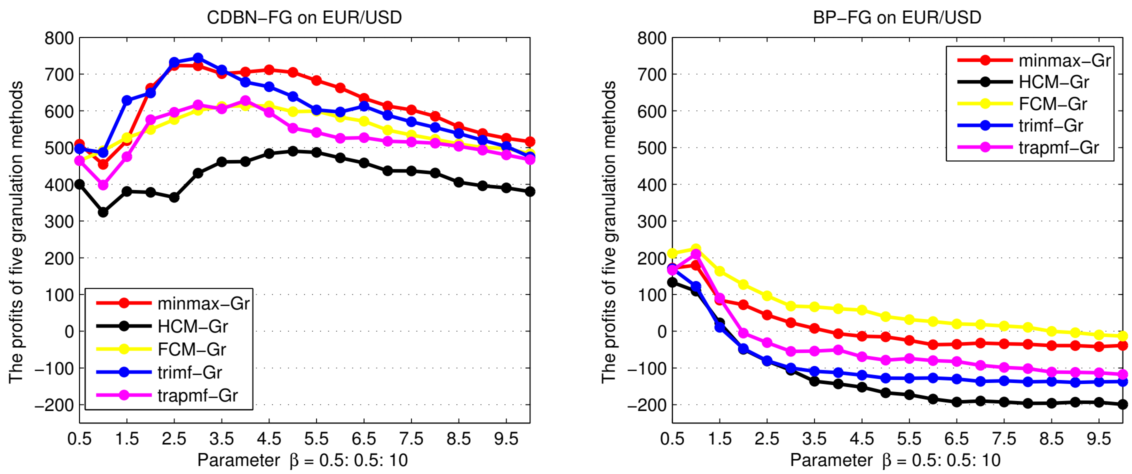

Fig. 2. The profits of CDBN-FG and BP-FG on EUR/USD

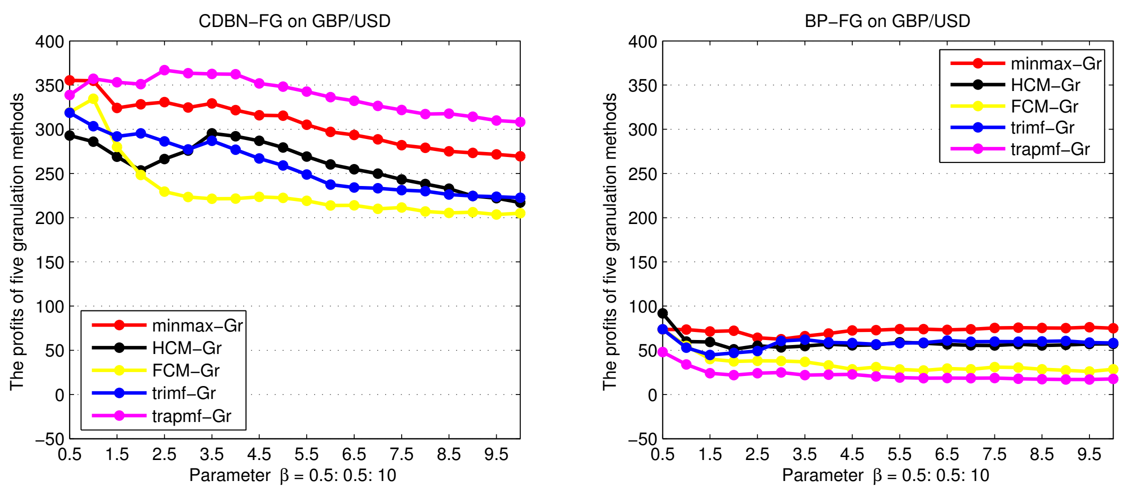

Fig. 3. The profits of CDBN-FG and BP-FG on GBP/USD

[18] L.A. Zadeh, “Fuzzy sets and information granulation”, in: M. Gupta, R.K. Ragade and R.R. Yager (Eds.), “Advances in Fuzzy Set Theory and Applications”, *North-Holland Publishing Company*, Amsterdam, New York, 1979, pp. 3-18.

[19] C. Mencar, “Theory of Fuzzy Information Granulation: Contributions to Interpretability Issues”, *Universit degli Studi di BARI*, Italy, 2004.

[20] H. Tahayori, W. Pedrycz and G.D. Antoni, “Distributed intervals: a formal framework for information granulation”, *Canadian Conference on Electrical and Computer Engineering*, pp. 1409-1412, 2007.

[21] H. Tahayori, A. Visconti and G.D. Antoni, “Email granulation based on augmented interval type-2 fuzzy set methodologies”, *IEEE International Conference on Granular Computing*, pp. 139-144, 2007.

[22] J.Y. Liang, J.H. Wang and Y.H. Qian, “A new measure of uncertainty based on knowledge granulation for rough sets”, *Information Sciences*, vol. 179, no. 4, pp. 458-470, 2009.

[23] Y.H. Qian, J.Y. Liang, Y.Y. Yao and C.Y. Dang, “MGRS: a multi-granulation rough set”, *Information Sciences*, vol. 180, no. 6, pp, 949-970, 2010.

[24] S.K. Meher and S.K. Pal, “Rough-wavelet granular space and classification of multispectral remote sensing image”, *Applied Soft Computing*, vol. 11, no.8, pp. 5662-5673, 2011.

[25] L.A. Zadeh, “Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic”, *Fuzzy Sets and Systems*, vol. 90, pp. 111-127, 1997.

[26] W. Pedrycz and G. Vukovich, “Feature analysis through information granulation andfuzzy sets”, *Pattern Recognition*, vol. 35, no. 4, pp. 825-834, 2002.

[27] S. Bandyopadhyay, S. Saha and W. Pedrycz, “Use of a fuzzy granulationCdegranulation criterion for assessing cluster validity”, *Fuzzy Sets and Systems*, vol. 170, no. 1, pp. 22-42, 2011.

[28] G.E. Hinton, “A fast learning algorithm for deep belief nets”, *Neural Computation*, vol. 18, pp.1527-1554, 2006.

[29] A.F. Murray, "Novelty detection using products of simple experts–a potential architecture for embedded systems", *Neural Networks*, vol. 14, pp. 1257-1264, 2001.
[30] H. Chen and A.F. Murray, "Continuous restricted Boltzmann machine with an implementable training algorithm", *IEE Proc-Vis. Image Signal Process*, vol. 120, pp. 153-158, 2003.