Chapter 6
Gold Price Forecasting Based on RBF Neural Network and Hybrid Fuzzy Clustering Algorithm

Fengyi Zhang and Zhigao Liao

Abstract This paper predicts good price based on RBF neural network employing hybrid fuzzy clustering algorithm. PCA technique has been used to integrate the 6 parameter dependent sub-variables of each TI (Technical Indicators, include MA, ROC, BIAS, D, K), which has been originated from the gold price before, and the results act as input. By employing a new hybrid fuzzy clustering algorithm, which is proposed by Antonios and George [10], *K*-Mean clustering algorithm and RBE algorithm, the predictions of price are yielded for each interval-*n* model. *n* refers to the number of predictions achieved by 1 operation. The most important conclusion indicates that the hybrid fuzzy clustering algorithm is superior to the general RBF central vector selecting algorithm mentioned above, in the aspects of MSE, *P*-Accuracy Rate and ROC.

Keywords Gold price forecasting · RBF neural network · PCA · Hybrid fuzzy clustering algorithm

6.1 Introduction

Forecasting gold price is becoming more and more important. For long in history, gold has been traded actively on international markets. Many derivatives of gold trading in international gold markets are also traded, such as gold futures, gold options, gold forward contracts, and so on [1, 2]. Remarkably, since the price of gold varies within a limited range, gold is able to reduce the effect of inflation, control the rise of price and help carry out constrictive monetary policy [3]. Hence, gold becomes an essential tool for risk hedging as well as an investment avenue. There-

F. Zhang · Z. Liao (✉)
Department of Management, Guangxi University of Science and Technology, Liuzhou 545000, P. R. China
e-mail: liaozhigao@126.com

J. Xu et al. (eds.), *Proceedings of the Seventh International Conference on Management Science and Engineering Management (Volume 1)*
Lecture Notes in Electrical Engineering 241, DOI: 10.1007/978-3-642-40078-0_6,

fore, to investors, it has become very significant and important to predict the price of gold.

The use of neural networks in forecasting the gold price has been operated before. McCann and Kalman [4] make an effort to use recursive neural networks to recognize the inflection points in the gold market based on historical data of ten indices, coming up with predictions that are both meaningful and profitable for the period studied. Tsibouris and Zeidenberg [5] and White [6] work with neural networks to forecast stock market indexes and individual assets. More recently, McMillan [7], using recursive and rolling estimation, find evidence of STAR nonlinearity being present within the DJIA. Further, the parameters of interest exhibit some temporal dependence. These results suggest that nonlinearity is a regular feature of the data that should be modeled and used in forecasting, although variations in parameter values may need to be incorporated. Chen and Leung [8] performance an evaluation of neural network architectures applied to the prediction of foreign exchange correlations, comparing the performance of models based on two competing neural network architectures, the multi-layered feed forward neural network (MLFN) and general regression neural network (GRNN). Their empirical evaluation measures the network models' strength on the prediction of currency exchange correlation with respect to a variety of statistical tests. The results of the experiments suggest that the selection of proper architectural design may contribute directly to the success in neural network forecasting. In addition, market-timing tests indicate that both MLFN and GRNN models have economically significant values in predicting the exchange rate correlation. Lai et al [9] propose a hybrid synergy model integrating exponential smoothing and neural network. 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.

The rest of this paper is organized as follows. Sect. 6.2 presents the traditional RBF neural network model and hybrid fuzzy clustering algorithm. In Sect. 6.3, the measurements of performance are thrown light on. Sect. 6.4 provides the experimental process and results by comparison. Finally, in Sect. 6.5, conclusions will be drawn.

6.2 Methodology

6.2.1 Traditional RBF Neural Network

The basic topology of the RBF network comprises in sequence a hidden layer and a linear processing unit forming the output layer. It is a kind of topology for a multi-input single-output network, where c represents the number of nodes in the hidden layer. Each hidden node corresponds to a radial basis function, while the output layer computes the weighted sum of the nodes' outputs. A radial basis function represents a local effect, the range of which is determined by its center element and width (variance). Herein, the radial basis function will also be referred to as kernel function or simply kernel. Employing the nomenclature of the topology mentioned above, the set of inputCoutput data pairs is symbolized as $S = \{(x_k, y_k) \in \Re^p \times \Re | f(x_k) = y_k, 1 \leq k \leq n\}$, n is the number of training samples, $x_k = [x_{k1}, x_{k2}, \cdots, x_{kp}]^T$ is the k-th input vector and y_k is the k-th output sample. We select Gaussian type kernel functions of the form:

$$g_i(x_k) = \exp\left(-\frac{\|x_k - v_i\|^2}{\sigma_i^2}\right), \tag{6.1}$$

where $v_1, v_2, \cdots, v_i, \cdots, v_c$ arise in the form of p-dimensional vectors and are referred to as kernel centers, and $\sigma_1, \sigma_2, \cdots, \sigma_i, \cdots, \sigma_c$ are the respective kernel widths.

Although too much effort has been put on calculating appropriate values for the kernels' centers, there are relatively few methods that address the issue of estimating the widths. Moody and Darken calculated each width using the average distance of the respective cluster center to its τ nearest neighbors,

$$\sigma_i = \frac{1}{\tau}\sqrt{\sum_{j=1}^{\tau} d_{ij}^2}, \tag{6.2}$$

where $d_{ij} = \|v_i - v_j\|$ with $i \neq j$, and typical values of τ are $\tau = 2$ and $\tau = 3$. A special case of Equation (6.2) was introduced by Pal and Bezdek [11] , where the width of each node was calculated by the distance between the center of the kernel and its nearest neighbor, multiplied by a positive factor.

6.2.2 Hybrid Fuzzy Clustering Algorithm

The K-Means algorithm is very sensitive to initialization but it is a fast procedure, while the fuzzy K-Means is able to reduce the dependence on initialization but it remains a slow process [11]. In a recent publication, Antonios and George [10] have developed a fuzzy learning vector quantization algorithm for image compression tasks, which combined the K-means and the fuzzy k-means. The basic idea of this paper is originated on that learning algorithm and utilizes the following objective function:

$$J_H = \theta \sum_{k=1}^{n}\sum_{i=1}^{c} u_{ik}\|x_k - v_i\|^2 + (1-\theta)\sum_{k=1}^{n}\sum_{i=1}^{c} (u_{ik})^2\|x_k - v_i\|^2, \tag{6.3}$$

where K is the number of clusters, $\theta \in [0, 1)$, and $u_{ik} \in [0, 1]$ is the membership degree of the k-th training vector to the i-th cluster. Notice that when $\theta = 0$, the objective function is transformed to the fuzzy k-means with $m = 2$, and when $\theta = 1$

it becomes the k-means algorithm. Therefore, the function posses a hybrid structure enabling the switch from fuzzy to crisp conditions depending on the value of θ.

Antonios and George [10] define the set T_k as the aggregate of the cluster centers affected by x_k. Initially, the set T_k includes all cluster centers and its cardinality is: $\aleph(T_k^{(0)}) = c$, where c is the number of radial basis. The proposed hybrid clustering algorithm can be operated as follows: Select values for c, and θ. Randomly initialize $v_1, v_2, \cdots, v_3$. Set $v = 0$, $\forall k : \aleph(T_k^{(0)}) = c$, and $T_k^{(0)} = \{v_1, v_2, \cdots, v_c\}$.
Step 1. Set $v = v + 1$.
Step 2. Update the sets $T_k^{(v)}$ and their cardinalities $\aleph(T_k^{(v)})(1 \leq k \leq n)$.
Step 3. Calculate the membership degrees $u_{ik}(1 \leq k \leq n; 1 \leq i \leq c)$.
Step 4. If $u_{ik} < 0(1 \leq k \leq n; 1 \leq i \leq c)$ then set $u_{ik} = 0$.
Step 5. Calculate the normalized membership degrees.
Step 6. Update the cluster centers.
Step 7. If there are no noticeable changes for the cluster centers then stop, else turn the algorithm to Step 1.

6.3 Measurement of Performance

(1) MSE

MSE is short for Mean Square Error, which measures the overall predicting ability of models. We define $r(s_i)$ as the real value of s_i and $p(s_i)$ the prediction of real value. Besides, $s_i \in S$ and $\mathrm{MSE}(S) = \sum(r(s_i) - p(s_i))^2/(\aleph(S) - 1)$ is the value of MSE over set S. Employing such a measurement, we can measure the degree of the deviation from real price to prediction.

(2) ROC

To measure the prediction performance, the area under the curve (AUC), which is defined as the area under the receiver operating characteristic (ROC) curve, is used. The ROC curve plots true positive rate as a function of false positive rate for differing classification thresholds. The AUC measures the overall quality of the ranking induced by model rather than the quality of a single value of threshold in that ranking. The closer the curve follows the left-hand border and then the top-border of the ROC space, the larger value of AUC the model produces, the more accurate the model is.

(3) P-Accuracy Rate

P-Accuracy Rate is a measurement used to detect how many predictions of value are close to the real value with p percent deviation. In this term, p refers to a variable number such as 0.1, 0.5 and 1. We define $r(s_i)$ as the real value of s_i and $p(s_i)$ the prediction of real value. Else, $A_p(S) = \{s_i \in S| \quad |r(s_i) - p(s_i)|/r(s_i) \leq p\%\}$ and $AR_p(S) = \aleph(A_p(S))/\aleph(S)$ is the value of P-Accuracy Rate over set S. By adjusting the value of P, an accuracy distribution is achieved, which plays an important role in comparison.

6.4 Experiment

There are 6 kinds of models: Interval-10, Interval-20, Interval-30, Interval-3, Interval-2 and Interval-1 in the Experiment. In Interval-*n*, *n* refers to the number of prediction of price achieved by 1 operation. This Experiment contains 3 parts, which are called data preprocessing, pretest and principal process, while the first and the second serve the third. In the first parts, 5 TIs (Technical Indicators, include MA, BIAS, ROC, K and D), which has been originated from the price before t, and the real price at the time of t are assembled as sample t. The samples set was separated for the reason that the two sample sets has been identified independent by statistic test. All samples was sorted by ascending time series. In the second part, we have used part of the pre-processed data and each group of parameters to forecast the gold price, and got the MSEs belonging to different groups of parameters. In the third part, rolling operation was employed to yield the result. 270 predictions in the big samples sets and 30 predictions in the small one were achieved. After calculating the MSE (Mean Square Error), CDPA (Correct Direction Predicted Rate), *P*-Accuracy Rate and ROC by the ways to compare the predicted price with the real price at the same time, to measure the performance of the models was not groundless. We have assessed the superiority of each algorithm by compare the MSE, CDPA, *P*-Accuracy Rate and ROC of the same Interval-*n* model.

6.4.1 Data Preprocessing

The data were downloaded from the website of WBG (World Bank Group), ranges from January-1975 to December-2011. The monthly averages of gold price were selected as our material for it could present the price lever of the every month better. There were 444 samples in total before data preprocessing.

The purpose of data preprocessing was to achieve the proper input which could supply enough information to forecast gold price, but also had the least dimensions.

The parameters for each TI were set as $P \in \{3,4,6,8,9,12\}$. Since a single price data was transformed to 5 TIs by the equation proposed in sheet n, and each of which has 6 dependent sub-variables, then the total number of the sub-variables per variable, or simply input dimensionality, became 30 (= 5TIs $\times$ 6 parameter dependent sub-variables). Even though the use of TI facilitates the consideration of the trends and the structure of the data, there was, on the other hand, the drawback that one variable has turned into a set of an increased number of sub-variables. The increased number of input variables means an increase in dimensionality, which degrades the performance of the prediction model. If we extract a single feature per TI, then the 6 parameter dependent sub-variables will be reduced to one dimensional feature. For PCA, this implies that we use only the first principal component from the co-variance matrix of the 6 parameter dependent sub-variables. After the processing above, we put the last 420 samples into our samples set, ranged from January-1977 to December-2011.

Subprime Crisis has led to the decrease of confidence for credit currency, but also became an important factor that pushed the gold price soaring [13]. Hence, to divide the previous set into two sets, which refer to the samples range from January-1977 to October-2007 before Subprime Crisis and the samples range from November-2007 to December-2011 after Subprime Crisis, is reasonable. Interval-10, Interval-20 and Interval-30 was used to forecast gold price in the big samples set which has 370 samples, while Interval-1, Interval-2 and Interval-3 are put into forecasting gold price in the small samples set which has 50 samples.

6.4.2 Pretest

The purpose of pretest was to find the seemingly proper group of parameters for each Model employing the hybrid fuzzy clustering algorithm. A consecutive area of data, whose size equals to the size that the Interval-*n* model requires, were employed to act as the 'historical knowledge' randomly for a certain model. For $n = 30, 20$ and 10, the corresponding size was 100. Likely, we has changed the size to be 20 for $n = 3, 2$ and 1. Then, each group of parameters and the 'historical knowledge' were employed by each model to get the forecasting price. And an n-width area next to the 'historical knowledge' was used to test the forecasting price.

Table 6.1 MSEs over pretest in Interval-10

MSE		$\theta = 0.3$	$\theta = 0.4$	$\theta = 0.5$	$\theta = 0.6$	$\theta = 0.7$
Iteration	4	19.653	2.491	3.593	111.344	140.709
	5	32.859	3.346	2.691	5.795	374.596
	6	9.866	3.819	1.102	48.692	2.0104
	7	6.359	4.062	4.991	27.360	2.887
	8	185.856	6.191	6.993	6.870	2.055
	9	174.705	48.852	51.045	0.501	29.223
	10	4.544	59.362	49.430	4.471	1.109
	11	5.523	31.385	59.319	4.707	1.431
	12	2.253	20.784	6.058	29.892	2.403
	13	1.790	4.086	2.812	16.309	43.483
	14	4.056	26.300	15.894	3.452	11.554
	15	2.699	3.368	3.536	2.176	27.584
	16	2.877	3.764	47.088	1.724	7.802

LMS (Least Mean Square) principle was used to measure the performance of the certain model with different group of parameters. The parameters we used for this experiment were composed of θ (which has a great effect on the function of the forecast model and have belonged to the set $\{0.3, 0.4, 0.5, 0.6, 0.7\}$) and the number of iterations (which refers to the times that the hybrid fuzzy clustering algorithm be operated to choose the proper radial basis. In this Experiment, it only included integers range from 4 to 16). To select the optimal θ-iteration group which related to the least MSE for a certain interval-n was of great importance. Such as, in Table 6.1 ($\theta = 0.6$, iteration $= 9$) was selected as the corresponding parameter group for the interval-10, for no other group has a less MSE. Likely, we have chosen ($\theta = 0.5$, iteration $= 6$), ($\theta = 0.6$, iteration $= 9$), ($\theta = 0.5$, iteration $= 6$), ($\theta = 0.7$, iteration $=$ 9) and ($\theta = 0.7$, iteration $= 6$) as the group of parameters for Interval-20, Interval-30, Interval-3, Interval-2 and Interval respectively.

6.4.3 Principal Process

The purpose of this part was to yield the prediction of gold price in each samples set. In addition, the algorithms of the models have contained hybrid fuzzy clustering algorithm, K-Mean clustering algorithm and the RBE algorithm which can perfectly fitting the multi-dimensions curve by employing m (m equals to the number of input-sample) radial basis.

While applying the models that disposing of the big samples set, we has used the first 100 samples, as 'historical knowledge', to yield the predicted price from 101^{th} to $100+n^{th}$. In interval-10, n refers to 10, and the like in other models. Assuming n equals to 10, as the process proceeding in a rolling operation, finally the prediction for the price from 361^{th} to 370^{th} could be yielded by using the samples from 261^{th} to 360^{th} as 'historical knowledge'. In operating small samples sets, excepted that the number of samples to be 'historical knowledge' was initialed by 20 and added to $20+i*n$ (i refers to the iteration that had been finished before) as the forecast process proceeding. Other processes are similar to the big one.

The difference between the two sorts of models lied in the size of samples set. The big one had adequate samples, while the small one has insufficient samples. By trial and err, it was verified that more 'historical knowledge' does not lead to more accurate result. On the contrary, it may bring in 'noise' which has decreased the prediction precise because of the 'out-of-date samples'. In the case of big samples set, 100 was a proper size for 'historical knowledge'. However, in the small one, the system was not stable for the small size of samples. To achieve a balance between high prediction precise and adequate predictions, the size of 'historical knowledge' was initialized by 20 and increases as the process proceeding.

The number of radial basis has also played an important role in our prediction process. Small-size radial basis model had poor ability to analyze, extract and restore the key information from the input, while too much radial basis have led to the 'over-approximating' situation, in which too much detail of the samples have been reserved however the intrinsic rule has not been recognized. By trial and error, 0.9 was a proper ratio between the number of radial basis and the size of input for hybrid fuzzy clustering algorithm and k-mean clustering algorithm.

6.4.4 Results of the Comparison: Between Algorithms

Abiding by the method mentioned above, 18 groups of prediction price, which refers to 6 kinds of models employing 3 different kinds of algorithm, have been achieved. After calculating the 4 kinds of performance measures, Interval-3 in the small set and Interval-10 in the big set were chosen to illustrate the excellent overall performance of hybrid fuzzy clustering algorithm, comparing with other algorithm, in Fig. 6.1, for they show the best performance in the aspect of *P*-Accuracy Rate, which is considered of great value.

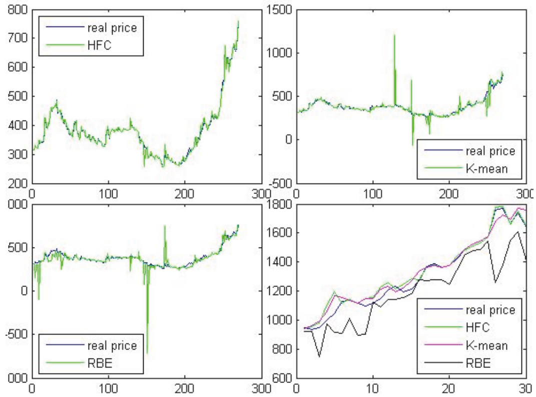

Fig. 6.1 Comparison between predictions and real price in Interval-10 and Interval-3

(1) MSE

As was mentioned above, MSE was a very important measurement for it can measure the degree how the predicted price deviates from the real price properly. In Table 6.2, the MSE-*m* referred to the MSE of the first *m* predictions.

Apparently, the results show that the hybrid fuzzy clustering algorithm always had a least MSE, compared to the *K*-Mean clustering and RBE, in each model. For example, in Interval-10, the MSE of hybrid fuzzy clustering was 34.230, seems to be much smaller than 459.280 and 6922.130, that of the *K*-Mean clustering and RBE. This phenomenon has appeared in each of model over the experiment. It means that the hybrid fuzzy clustering algorithm is superior to the *K*-Mean clustering and RBE in forecasting.

Another interesting thing was that, the MSE-m of hybrid fuzzy clustering and *K*-Mean clustering in each model had a trend of decreasing with respect to the increasing of m. It means that, as the size of 'history knowledge' increasing, the forecast ability of hybrid fuzzy clustering and that of *K*-Mean clustering is strengthening.

(2) ROC

As was mentioned above, ROC is a very important measurement for it can measure the ability to predict the direction of the price. Interval-1 in the small set and Interval-10 in the big set are chosen to illustrate the excellent overall performance of hybrid fuzzy clustering algorithms, comparing with other algorithm in Table 6.3,

Table 6.2 MSEs in each modes

MSE		Algorithm		
		HFC	*K*-Mean	RBE
Interval-1	MSE-5	4439.996	4635.752	7120.005
	MSE-10	2357.829	2397.470	8172.744
	MSE-15	1940.494	2109.092	5895.572
	MSE-20	1741.743	1794.250	5108.762
	MSE-25	1379.549	1423.007	4146.090
	MSE-30	1429.305	1443.070	13016.080
Interval-2	MSE-5	2933.092	5211.516	11619.470
	MSE-10	1495.537	2598.631	14507.470
	MSE-15	1593.897	2256.629	10168.720
	MSE-20	1433.599	1856.365	9770.745
	MSE-25	1135.340	1470.952	7906.600
	MSE-30	1289.372	1547.582	19533.740
Interval-3	MSE-5	8888.387	5289.069	14447.270
	MSE-10	4513.322	2864.558	23127.380
	MSE-15	4066.433	2518.297	16186.820
	MSE-20	3005.126	2109.878	14325.570
	MSE-25	2388.210	1683.368	11830.000
	MSE-30	1999.735	2242.144	26469.620
Interval-10	MSE-270	34.230	459.280	6922.130
Interval-20	MSE-270	256.370	28415.330	28415.330
Interval-30	MSE-270	110.480	9385.840	10643.640

Table 6.3 AUCs in each modes

AUC	HFC	*K*-Mean	RBE
Interval-1	0.956	0.955	0.818
Interval-2	0.935	0.948	0.773
Interval-3	0.948	0.890	0.786
Interval-10	0.984	0.843	0.734
Interval-20	0.951	0.823	0.679
Interval-30	0.970	0.782	0.653

for they show the best performance in the aspect of AUC, which refers to the area under curve and be considered as the principal character of ROC.

Apparently, the results have showed that the hybrid fuzzy clustering algorithm always has a best ROC, compared to the *k*-mean clustering and RBE, in each model. For example, in Interval-10, the AUC of hybrid fuzzy clustering is 0.984 seems to be much better than 0.843 and 0.734, which are the value of the *K*-Mean clustering and RBE. This phenomenon appears in each of model over the experiment. It means that the hybrid fuzzy clustering algorithm is superior to the *K*-Mean clustering and RBE in forecasting.

(3) *P*-Accuracy Rate

As was mentioned above, P-Accuracy Rate was a very important measurement for it can measure the degree how the predicted price deviated from the real price properly. In Table 6.4, there were 3 kinds of value of P, belonged to the set $\{1, 0.5, 0.1\}$.

Table 6.4 P-Accuracy Rates in each modes

P-Accuracy Rate		$P = 1$	$P = 0.5$	$P = 0.1$
Interval-1	HFC	0.433	0.333	0.100
	K-Mean	0.467	0.400	0.133
	RBE	0.300	0.167	0.067
Interval-2	HFC	0.500	0.400	0.167
	K-Mean	0.500	0.400	0.100
	RBE	0.100	0.100	0.067
Interval-3	HFC	0.600	0.300	0.133
	K-Mean	0.367	0.233	0.033
	RBE	0.067	0.000	0.000
Interval-10	HFC	0.870	0.740	0.220
	K-Mean	0.628	0.468	0.063
	RBE	0.3978	0.2602	0.0706
Interval-20	HFC	0.848	0.744	0.249
	K-Mean	0.599	0.450	0.078
	RBE	0.364	0.260	0.082
Interval-30	HFC	0.836	0.673	0.234
	K-Mean	0.595	0.454	0.138
	RBE	0.338	0.223	0.052

Apparently, the results have showed that the hybrid fuzzy clustering algorithm always had a best P-Accuracy Rate, compared to the K-Mean clustering and RBE, in each model. For example, in Interval-10, the P-Accuracy Rate of hybrid fuzzy clustering was 0.870, seems to be much better than 0.628 and 0.398, that of the K-Mean clustering and RBE. This phenomenon has appeared in each of the models over the experiment. It means that the hybrid fuzzy clustering algorithm is superior to the K-Mean clustering and RBE in forecasting.

6.4.5 Results of the Comparison: Between Models

The issue we should focus on is that employing hybrid fuzzy clustering algorithm, the Interval-3 performs better than the Interval-2 colliding with the assumption that the short-interval models should be superior to that of the long-interval models. The reason why assume above condition was that as the interval increasing, the prediction would originate from the more upgraded set, which has more important information for the certain prediction. Considering such an issue, there is a vital difference between the Interval-2 and the Interval-3. With the different group of parameters, the small interval performs better than the big one becomes not sure.

Table 6.5 Comparison of interval-2 and interval-3 employing hybrid fuzzy clustering with the same parameters group ($\theta = 0.5$, iteration = 6)

Value		Interval	
		2	3
MSE	MSE-5	3431.386	8888.387
	MSE-10	2441.681	4513.322
	MSE-15	2432.076	4066.433
	MSE-20	2146.086	3005.126
	MSE-25	1758.642	2388.210
	MSE-30	2863.197	1999.735
P-Accuracy Rate	$P = 1$	0.333	0.600
	$P = 0.5$	0.200	0.300
	$P = 0.1$	0.067	0.133
ROC	AUC	0.916	0.948

So, to set the same group of parameters to the Interval-2 and the Interval-3 is advisable. By recalculating the predicted price, MSE, *P*-Accuracy Rate, CDPR and ROC, the final results were presented in Table 6.5, and the parameters of Interval-2 and Interval-3 were both ($\theta = 0.5$, iteration = 6). In this way, the assumption was testified to be untrue. For applying the same parameters group, the Interval-2 was still inferior to the Interval-3 in all of the aspects.

6.5 Conclusions

There are 4 conclusions can be drawn from the experiment:

- Hybrid fuzzy clustering algorithm is superior to *K*-Mean clustering and RBE on the ability to generalize.
- The generalizing ability of hybrid fuzzy algorithm increase with respect to size of 'historical knowledge'.
- The parameters group for hybrid fuzzy clustering algorithm effects the performance deeply.
- While employing hybrid fuzzy clustering algorithm, the short-interval models may not generate more precise results compared to the long-interval models.

References

1. Grudnitski G, Osburn L (1988) Forecasting S&P and gold futures prices: An application of neural networks. Journal of Futures Markets 13(6):631–643
2. Shafieea S, Topalb E (2010) An overview of global gold market and gold price forecasting. Resources Policy 35(3):178–189
3. Atsalakisa GS, Valavanisb KP (2009) Suveying stock market forecasting techniques — Part II: Soft computing methods. Expert Systems with Applications 36(3):5932–5941
4. McCann P, Kalman B (1994) A neural network model for the gold market. Washington University in St. Louis, Part of the Philosophy Neuroscience and Psychological Technical Report Series
5. Tsibouris G, Zeidenberg M (1995) Testing the efficient markets hypothesis with gradient descent algorithms. Neural Networks in the Capital Markets 127–136
6. White H (1993) Economic prediction using neural networks: The case of IBM daily stock returns. Neural Networks in Finance and Investing. Probus Publishing 315–328
7. McMillan D (2005) Non-linear dynamics in international stock market returns. Review of Financial Economics 14:81–91
8. Chen A, Leung MT (2005) Performance evaluation of neural network architectures: The case of predicting foreign exchange correlations. Journal of Forecasting 24(6):403–420
9. Lai KK, Yu L, Wang S, Huang W (2006) Hybridizing exponential smoothing and neural network for financial time series predication. Lecture Notes in Computer Science. Springer Berlin/Heidelberg
10. Antonios DN, George ET (2012) A novel training algorithm for RBF neural network using a hybrid fuzzy clustering approach. Fuzzy Sets and Systems 193:62–84
11. Pal NR, Bezdek JC (1995) On clustering validity for the fuzzy *c*-means model, IEEE Transactions on Fuzzy Sets and Systems 3:370–379
12. Saha A, Keeler JD (1989) Algorithms for better representation and faster learning in radial basis function networks. Advances in Neural Information Processing Systems 482–489
13. Zhang Y (2011) Analysis of gold price and its effect factors. University of Zhejiang, Hangzhou (In Chinese)