The application of neural networks to forecast fuzzy time series

Amaal El Sayed Abd El Ghany Mubarak
Department of applied statics and insurance,
Faculty of commerce, Damietta University, Egypt

Abstract

This study applies a back-propagation neural network to forecast fuzzy time series. Three models are proposed; a conventional fuzzy time series model and two hybrid models. Hybrid1 model uses a neural network approach to establish fuzzy relationships in fuzzy time series and hybrid2 model uses a neural network approach to improve forecasts from the conventional fuzzy time series model. The daily prices of golden pound for October 2014 were chosen as the forecasting target. The empirical results show that the hybrid2 model outperforms both the conventional fuzzy time series and the hybrid1 models.

Keywords: *Back-propagation; Forecasting; Golden pound; Fuzzy time series.*

1. Introduction

Numerous fuzzy time series models have been proposed in scientific literature during the past decades. Among the most accurate fuzzy time series models found in literature are the high order models. However the current prediction methods have not been able to provide satisfactory accuracy rates for defuzzified outputs (forecasts). This study, in addition to showing how to apply the conventional fuzzy time series model, it also shows how neural networks can be used to establish fuzzy relationships and to improve forecasts. A fuzzy time series essentially consists of steps such as fuzzification, the establishing of fuzzy relationships, and defuzzification. This study has chosen a neural network to establish fuzzy relationships in fuzzy time series, which are also nonlinear. Hence, the rational and motivation for applying a neural network are self-explanatory. We have followed the suggestions regarding constructing a neural network for forecasting, including data preparation, the network setup, and model selection and evaluation. To investigate the forecasting capabilities of the neural network approach, we propose three models: a conventional fuzzy time series model, a hybrid1 model uses a neural network approach to establish fuzzy relationships and a hybrid2 model uses a neural network approach to improve forecasts from the conventional fuzzy time series model. The forecasting results of both hybrid models are then compared with those from the conventional fuzzy time series model. To show these things, the remainder of this study is organized as follows. Section 2 deals with relevant theoretical aspects of fuzzy time series. Section 3 describes the setup for the hybrid models using the neural network approach. Section 4 explains the empirical results. Finally, section 5 concludes the study.

2. Fuzzy time series

Let U be the universe of discourse, where $U = \{u_1, u_2, \ldots, u_b\}$. A fuzzy set A_i of U is defined as $A_i = f_{A_i}(u_1)/u_1 + f_{A_i}(u_2)/u_2 + \ldots + f_{A_i}(u_b)/u_b$,where f_{A_i} is the membership function of the fuzzy set A_i; $f_{Ai}: U \to [0,1]$. u_a is an element of fuzzy set A_i; $f_{A_i}(u_a)$ is the degree of belongingness of u_a to A_i; $f_{A_i}(u_a) \in [0,1]$ and $1 \le a \le b$.

Definition [1] $Y(t)(t = \ldots, 0, 1, 2, \ldots)$ is a subset of a real number. Let $Y(t)$ be the universe of discourse defined by the fuzzy set $f_i(t)$. If $F(t)$ consists of $f_i(t)(i = 1, 2, \ldots)$, $F(t)$ is defined as a fuzzy time series on $Y(t)(t = \ldots, 0, 1, 2, \ldots)$. Following Definition 1, fuzzy relationships between two consecutive observations can be defined as follows:

Definition [2] If there exists a fuzzy relationship $R(t-1, t)$, such that $F(t) = F(t-1) \times R(t-1, t)$, where $\times$ represents an operation, then $F(t)$ is said to be caused by $F(t-1)$.

Definition [3] Let $F(t-1) = A_i$ and $F(t) = A_j$. The relationship between two consecutive observations, $F(t)$ and $F(t-1)$, referred to as a fuzzy logical relationship (FLR), can be denoted by $A_i \to A_j$, where A_i is called the left- hand side and A_j the right- hand side of the FLR. Song and Chissom (1993) proposed a fuzzy time series model, that include the following steps. Step 1: define and partition the universe of discourse. Step 2: define fuzzy sets for the observations. Step 3: fuzzify the observations. Step 4: establish the fuzzy relationship, *R*. Step 5: forecast and step 6: defuzzify the forecasting results. Among these steps, the establishing of the fuzzy relationships directly affects the forecasting results; hence, this has become the target of many relevant studies. Chen (1996) proposed a model based on arithmetic operations. The fuzzy relationship was established by putting the same LHS of the FLRs together into fuzzy logical relationship groups (FLRGs). For example, there are FLRs with the same LHSs $(A_i): A_i \to A_{j1}$, $A_i \to A_{j2}, \ldots$ these FLRs can be grouped into an FLRG as $A_i \to A_{j1}, A_{j2}, \ldots$

3. Neural network-based fuzzy time series models

A forecasting neural network model consists of the following steps, namely, data preparation, the neural network setup (input variable selection, the choice of structure, the transfer function, etc.), and evaluation and selection. In the setup, we intended to establish (or train) the fuzzy relationships of all the FLRs and then to forecast. As in Definition 3, an FLR is a one to one relationship. Hence, there is one input layer and one output layer with one node each. Most forecasting applications such as K. Hornik (1993) have used only one hidden layer and a sufficiently large number of hidden nodes. Meanwhile, to prevent over-fitting, a small neural network was preferred. Accordingly, we used one hidden layer and two hidden nodes. Hence, we set up a neural network structure as in Fig. 1.

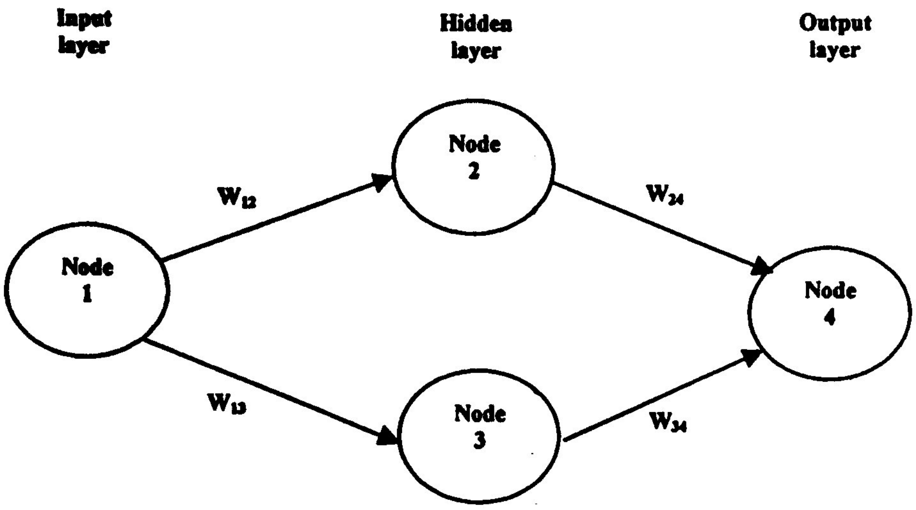

Fig. 1 . Neural network structure.

There is (are) input(s) to the node s from the node(s) r of the previous layer, such as X_r. Each connection from node r to s is associated with a weight W_{rs}, representing the connection strength in between. The output of node s; Y_s, is computed as follows:

$$Y_s = f(\sum W_{rs} \times X_r - \theta_r), \qquad (1)$$

$$f(Z) = \frac{1}{1+e^{-Z}} \qquad (2)$$

where $f(Z)$ is a sigmoid function.

We can then illustrate the proposed hybrid1 approach as in the following steps.

Step 1: Defining and partitioning the universe of discourse.

The universe of discourse for observations; U, is defined according to the problem domain, so U = [starting, ending]. After the length of intervals, l, is determined, the U can be partitioned into equal-length intervals $u_1, u_2, u_3, \ldots, u_b, b = 1, \ldots$ and their corresponding midpoints $m_1, m_2, m_3, \ldots, m_b$ respectively.

Step 2: Defining fuzzy sets for observations.

Each linguistic observation, A_i, can be defined by the intervals $u_1, u_2, u_3, \ldots, u_b$. $A_i = f_{A_i}(u_1)/u_1 + f_{A_i}(u_2)/u_2 + \ldots + f_{A_i}(u_b)/u_b$.

Step 3: fuzzifying the observations.

Each observation is fuzzified to a fuzzy set. In this context, fuzzification is the process of identifying associations between the historical values in the data set and the fuzzy sets defined in the previous step. As in Song and Chissom(1993), an observation is fuzzified to A_i if the maximal degree of membership of that observation is in A_i. If the highest degree of belongingness of a certain historical time variable, say $F(t-1)$, occurs at fuzzy set A_i, then $F(t-1)$ is fuzzified as A_i.

Step 4: Establishing the fuzzy relationship (neural network training).

We used the back-propagation neural network to establish the fuzzy relationships in these FLRs. For FLR, $A_i \rightarrow A_j$, i became the input and j its corresponding output.

Step 5: Forecasting.

In the hybrid1 model we suppose that $F(t-1) = A_{i'}$. We set i' as the input for forecasting. If the output from the neural network is j' then we say that the fuzzy forecast is $A_{j'}$. That is,

$$F(t) = A_{j'}. \qquad (3)$$

Step 6: Defuzzifying.

The defuzzified forecast is equal to the midpoint of the fuzzy forecast. Suppose the fuzzy forecast of $F(t)$ is $A_{k'}$. The defuzzified forecast is equal to the midpoint of $A_{k'}$, i.e., $forecast_t = m_{k'}$.

In order to make comparisons that show the best way to use neural networks in building a hybrid model to predict the daily price of golden pound ,we introduced the hybrid2 model. Hybrid2 model uses the neural networks approach to improve forecasts from the conventional fuzzy time series model. We consider the forecasts from the conventional fuzzy time series model as inputs in the neural network to get outputs that will be improved predictions.

4. Empirical analysis[1]

We used the daily price of golden pound during the period from January 2014 to October 2014 for our Empirical analysis. The data from January 2014 to September 2014 were used for estimation (training), while those for October 2014 were used for forecasting. We divided this Empirical analysis to four stages as follows:

stage1: (starting with the fuzzy time series model)

This stage includes the steps from 1 to 3 in the fuzzy time series model. The universe of discourse for observations; U, was defined as [2056.5, 2518.32]. The length of intervals was determined as 8.2468; U was partitioned into equal-length intervals from u_1 to u_{56}. The midpoints of these intervals were from m_1 to m_{56}, respectively. The intervals were set as $u_1 = [2056.5, 2064.7468], u_2 = [2064.7468, 2072.9936], \ldots$ and the midpoints were set as $m_1 = 2060.6234, m_2 = 2068.8702$.

1. The software used were matlab and spss

Table 1
The observation intervals

u_1=[2056.5 , 2064.7468]	U_{13}= [2155.4614 , 2163.7082]
u_2= [2064.7468 , 2072.9936]	U_{14}= [2163.7082 , 2171.955]
u_3= [2072.9936 , 2081.2404]	U_{15}= [2171.955 , 2180.2018]
u_4= [2081.2404 , 2089.4871]	U_{16}= [2180.2018 , 2188.4486]
u_5= [2089.4871 , 2097.7339]	U_{17}= [2188.4486 , 2196.6954]
u_6= [2097.7339 , 2105.9807]	U_{18}= [2196.6954 , 2204.9421]
u_7= [2105.9807 , 2114.2275]	U_{19}= [2204.9421 , 2213.1889]
u_8= [2114.2275 , 2122.4743]	U_{20}= [2213.1889 , 2221.4357]
u_9= [2122.4743 , 2130.7211]	U_{21}= [2221.4357 , 2229.6825]
u_{10}= [2130.721 , 2138.9679]	U_{22}= [2229.6825 , 2237.9293]
u_{11}= [2138.9679 , 2147.2146]	U_{23}= [2237.9293 , 2246.1761]
u_{12}= [2147.2146 , 2155.4614]	U_{24}= [2246.1761 , 2254.4229]
	....

Each linguistic observation A_i was defined as follows:

$$A_1 = 1.0/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + ... + 0/u_{55} + 0/u_{56}$$

$$A_2 = 0/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + ... + 0/u_{55} + 0/u_{56}$$

$$\vdots$$

$$A_{56} = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + ... + 0.5/u_{55} + 1.0/u_{56}$$

Each observation was fuzzified to a fuzzy set according to the maximal degree of membership of that observation. Some of the results of this step were as in Table 2.

Table 2
Fuzzy golden pound price

Date	golden pound price	Fuzzy golden pound price
1-Jan	2151.13	A12
2-Jan	2186.87	A16
3-Jan	2214.51	A20
4-Jan	2214.51	A20
5-Jan	2208.67	A19
6-Jan	2215.29	A20
7-Jan	2197.02	A18
8-Jan	2188.02	A16
9-Jan	2198.12	A18
10-Jan	2228.7	A21
11-Jan	2232.64	A22
12-Jan	2232.55	A22
13-Jan	2237.11	A22
14-Jan	2223.16	A21
15-Jan	2220.41	A20
16-Jan	2221.49	A21
17-Jan	2242.23	A23
18-Jan	2242.23	A23

Continue

19-Jan	2249.81	A24
20-Jan	2245.58	A23
21-Jan	2222.44	A21
22-Jan	2212.1	A19
23-Jan	2258.83	A25
24-Jan	2268.39	A26
25-Jan	2270.02	A26
26-Jan	2276.4	A27
27-Jan	2247.26	A24
28-Jan	2239.54	A23
29-Jan	2265.7	A26
30-Jan	2224.81	A21
...	...	...

Based on the Fuzzy sets in step 3, some FLRs were established as follows:

$A12 \rightarrow A16$, $A16 \rightarrow A20$, $A20 \rightarrow A20$, $A20 \rightarrow A19$ $A19 \rightarrow A20$, ...

Then we established fuzzy logical relationship groups (FLRGs), some of these were as follows:

Table 3

Fuzzy logical relationship groups

A12 ⟶	A16
A16 ⟶	A20, A18
A20 ⟶	A20, A19, A18, A21
A19 ⟶	A20, A25
A18 ⟶	A16, A21
A21 ⟶	A22, A20, A23, A19, A21, A21, A21, A24
A22 ⟶	A22, A22, A21, A24, A22, A23, A23, A24, A22, A20
A23 ⟶	A23, A24, A21, A26, A24, A26, A22, A22, A23, A22
.....	

We can then forecast, for example, by F (6) using the fuzzy time series (Chen's model) as the average of the midpoints of the intervals u_{20} and u_{25} since F (5) have the fuzzy set A19. The intervals $u_{20} = [2213.1889, 2221.4357], u_{25} = [2254.4229, 2262.6696]$ with midpoints equal 2217.3123and 2258.54625 respectively, so the forecast of F (6) equal 2237.92927.

Stage 2: (Using the back-propagation neural network to establish the fuzzy relationships (hybrid1 model))

Fuzzy golden pound price obtained from *stage 1* became the input and output patterns for neural network training. The fuzzy set at 1-January was considered as an input to the fuzzy set at 2-January and so on. The data from January to September were used for training (the in-sample) and October for forecasting (the out-of-sample). In other words, the ratio is about 90%:10%. One hidden layer was set up, the learning rate was initially set at 0.2, and the momentum was set at 0.8. The output layer had one neuron, which was the forecast value, i.e., the number of the fuzzy set. Some of the results are listed in Table 4.

Table 4
Fuzzy golden pound price from neural network training

Date	INPUT	OUTPUT
2-Jan	12	١٥
3-Jan	16	١٨
4-Jan	20	٢٠
5-Jan	20	٢٠
6-Jan	19	٢٠
7-Jan	20	٢٠
8-Jan	18	١٩
9-Jan	16	١٨
10-Jan	18	١٩
11-Jan	21	٢١
12-Jan	22	٢٢
13-Jan	22	٢٢
14-Jan	22	٢٢
15-Jan	21	٢١
...	...	...

In order to Forecasting and Defuzzifying in this stage we first used the forecasts of the fuzzy sets. For example, the golden pound price for 8 January was 2188.02, mapped to A_{16}. In other words, $F(t-1)=F(8)=A_{16}$ which corresponding to the input, 16, for 9 January and the output from the neural network was 18. That is $F(t)=F(9)=A_{18}$. Second, we defuzzifying the forecast. The defuzzified forecast is equal to the midpoint of the fuzzy interval. For example, $F(9)=A_{18};\textit{the forecast} \ =m_{18}=$ ٢٢٠٠.٨١٨٧٥.

Stage 3: (Using the back-propagation neural network to improvement the Forecasts from Chen's model (hybrid2 model))

In the hybrid2 model, we treated the obtained forecasts from Chen's model as inputs in neural network model. The neural network was trained using Chen's outputs ($\hat{z}_t, a_t$).

There were two neurons in the input layer; the forecast value $\hat{z}_t$, and residual value a_t. The neural network structure is characterized by a network of three layers of simple processing units connected by acyclic links (Fig. 2).

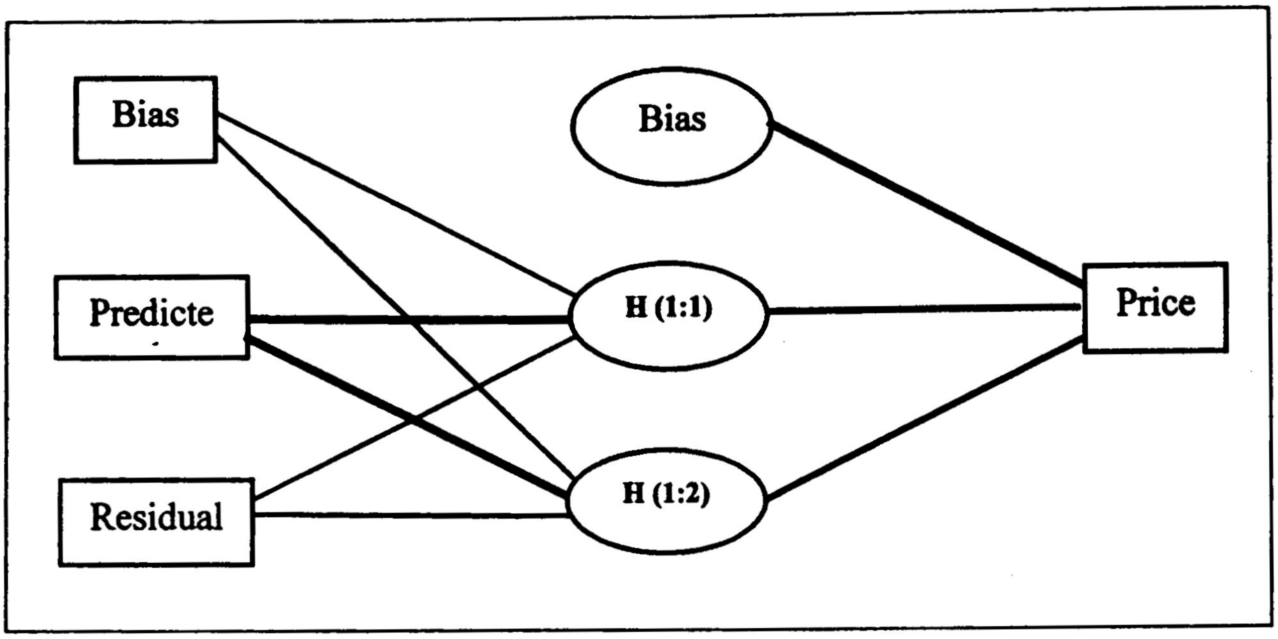

Fig. 2. Structure of the best fitted network

The forecasting results from these two hybrid models were then compared with those from a fuzzy time series model, Chen's model. For evaluation purposes, the forecast error and root mean squared error (RMSE) were used to measure performance:

$$error_t = |actual_t - forecast_t|,$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{c} \text{error}_t^2}{c}},$$

Where there are c forecasts.

The obtained outputs from the hybrid2 model was actually improved forecasts as well. The comparison between the forecasts obtained from the conventional fuzzy time series model; Chen's model, and the two proposed hybrid models show that the proposed two hybrid models outperform the conventional fuzzy time series model; Chen's model. Table 6 shows these results.

Table 5
The forecasts from the proposed models

Hybrid 2 model	Hybrid 1 model		Chen's model			
forecasts	forecasts	fuzzy sets	forecast	fuzzy set	golden pound price	Date
2228.63	2217.312	٢٠	2213.19	A21	2228.21	1-Oct
2230.7	2225.559	٢١	2229.68	A22	2230.89	2-Oct
2193.71	2233.806	٢٢	2233.81	A16	2187.77	3-Oct
2194.22	2200.819	١٨	2209.07	A16	2187.77	4-Oct
2203.13	2200.819	١٨	2209.07	A18	2198.84	5-Oct
2217.73	2209.066	١٩	2204.94	A20	2215.87	6-Oct
2224.81	2217.312	٢٠	2213.19	A21	2224.02	7-Oct
2246.23	2225.559	٢١	2229.68	A24	2247.28	8-Oct
2246.49	2250.3	٢٤	2247.55	A24	2248.02	9-Oct
2246.13	2250.3	٢٤	2247.55	A24	2247.65	10-Oct
2246.13	2250.3	٢٤	2247.55	A24	2247.65	11-Oct
2264.55	2250.3	٢٤	2247.55	A26	2266.21	12-Oct
2268.16	2266.793	٢٦	2250.3	A26	2269.84	13-Oct
2256.6	2266.793	٢٦	2250.3	A25	2258.37	14-Oct
2276.26	2258.546	٢٥	2258.55	A27	2277.93	15-Oct
2272.96	2275.04	٢٧	2269.54	A27	2275.02	16-Oct
2272.96	2275.04	٢٧	2269.54	A27	2275.02	17-Oct
2272.87	2275.04	٢٧	2269.54	A27	2274.93	18-Oct
2271.11	2275.04	٢٧	2269.54	A27	2273.22	19-Oct
2287.7	2275.04	٢٧	2269.54	A29	2289.17	20-Oct
2290.18	2291.533	٢٩	2291.53	A29	2292.10	21-Oct
2278.89	2291.533	٢٩	2291.53	A28	2281.32	22-Oct
2258.7	2283.287	٢٨	2291.53	A25	2261.55	23-Oct
2259.17	2258.546	٢٥	2258.55	A25	2261.16	24-Oct
2259.17	2258.546	٢٥	2258.55	A25	2261.16	25-Oct
2252.79	2258.546	٢٥	2258.55	A25	2254.74	26-Oct
2251.01	2258.546	٢٥	2258.55	A24	2252.93	27-Oct
2255.51	2250.3	٢٤	2247.55	A25	2257.20	28-Oct
2225.64	2258.546	٢٥	2258.55	A21	2226.12	29-Oct
2207.53	2225.559	٢١	2229.68	A18	2204.67	30-Oct
2171.25	2209.066	١٩	2204.94	A13	2156.23	31-Oct

Table 6
Comparison of models performance

The proposed model	RMSE
Chen's model	18.243
hybird model 1	16.94
hybird model 2	3.631

The forecasts from all of the proposed models are depicted in Fig. 3

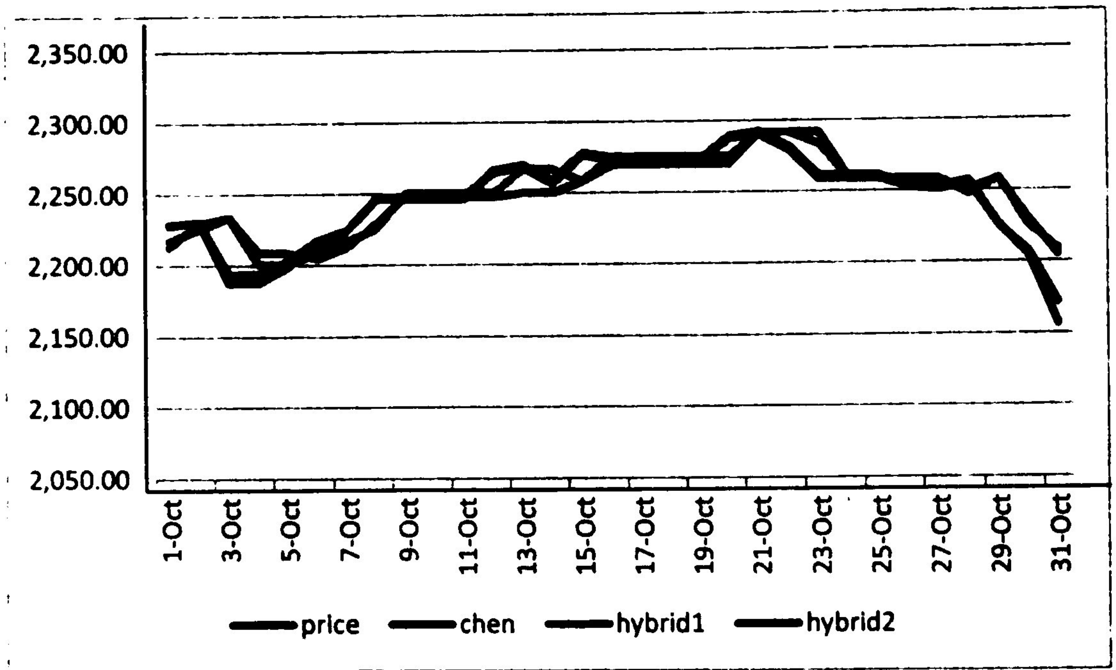

Fig. 3 Comparison of forecasts for October

5. Conclusions

This study has applied a back-propagation neural network for two purposes; first, to assist in fuzzy time series modeling, second, to improve forecasts. We have compared the forecasting results with a conventional fuzzy time series model, i.e. Chen's model to demonstrate the superiority of the proposed models. Actually, we have proposed two hybrid models for forecasting. The first model, hybrid1 model, applied neural networks to establish the fuzzy relationships and so we made Forecasting and Defuzzifying as *stage 2*. The other hybrid model, hybrid2, treated the obtained forecasts and residuals from Chen's model as inputs in a neural network to get improved forecasts. As a result, the hybrid2 model outperformed the hybrid1 model as well as the conventional fuzzy time series, Chen's model.

References

[1] Ansuj, A.P., Camargo, M.E., Radharamanan, R., and Petry, D.G. (1996). Sales forecasting using time series and neural networks, *Comput. Ind. Eng.* 31 (1): 421-424.

[2] Cabarkapa, S., Kojic, N., Savic, A., and Zivkoviv, B. (2010). Use of artificial neural networks in financial time series prediction and financial risk prediction, *Infoteh-Jahorina*,vol.q,Ref.E-III-6,P.584-587

[3] Caire, D., Hatabian, G., and Muller, C. (1992). Progress in forecasting by neural networks, *Int. Conf. Neural Networks II*, 540-545.

[4] Chen, S. M. (1996). Forecasting enrollments based on fuzzy time series, *Fuzzy Sets and Systems*, (81): 311-319.

[5] Chen, S. M. (2000). Forecasting enrollments based on high-order fuzzy time series, *Cyberent. Syst.:Int. J.* 33 (1):1-16.

[6] Hill, T., O'Connor, W., and Remus, M. (1996). Neural network models for time series forecasts, *Manage. Sci.* 42 (7): 1082-1092.

[7] Hornik, K. (1993). some new results on neural network approximation, *Neural Networks* (6):1069-1072.

[8] Huarng, K., and Yu, H.-k. (2004). A dynamic approach to adjusting lengths of intervals in fuzzy time series forecasting, Intell. Data Anal. 8 (1): 3-27.

[9] Huarng, K., and Yu, H.-k. (2006). The application of neural networks to forecast fuzzy time series, *Physica A* (363): 481-491.

[10] Hwang, J.R., Chen, S.M., and Lee, C.H. (1998). Handling forecasting problems using fuzzy time series, *fuzzy Sets and Systems*, (100): 217-228.

[11] Khashei, M., and Bijari, M. (2010). An artificial neural network (p, d, q) model for time series forecasting", *Expert Systems with Applications*, (37): 479-489.

[12] Khashei, M., and Bijari, M. (2011). Exchange rate forecasting better with hybrid artificial neural networks models, *J. Math. Comput. Sci.*, 1 (1): 103-125.

[13] Maier, H.R., Dandy, G.C. (1996). Neural network models for forecasting univariate time series, *Neural Networks World*, 6 (5): 747-772.

[14] Mubarak, A. (2014). Improving time series forecasting using a hybrid SARIMA and neural network model, *The Egyptian statistical journal*, 2 (58):118-133

[15] Song, Q., and Chissom, B.S. (1993). Forecasting enrollments with fuzzy time series – part I, *Fuzzy Sets and Systems*, (54):1-9.

[16] Yadav, V., Krishnan, R., Kumar, N., and Bharti, V. (2013). A comparative study of neural-network & fuzzy time series forecasting techniques – Case study: Marine fish production forecasting, *Indian Journal of Geo-marine Sciences*, 42(6):707-716.

[17] Zhang, B., Zhang, S., and Lu, G. (2013). Evaluation model research of 100 meters sprint exercise capacity based on fuzzy neural network, *Journal of chemical and pharmaceutical research*, 5 (9):256-262.